DOCUMENT RESUME

ED 116 807 95 PS 008 290

AUTHOR Conrad, Eva E.
TITLE The Effects of Tutor Achievement Level, Reinforcement Training, and Expectancy on Peer Tutoring.
INSTITUTION Arizona Univ., Tucson. Arizona Center for Educational Research and Development.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE 18 Jul 75
NOTE 76p.; Not available in hard copy due to marginal reproducibility of original document

EDRS PRICE MF-$0.76 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS *Academic Achievement; *Expectation; Feedback; Grade 2; High Achievers; Low Achievers; *Peer Teaching; *Primary Education; Reinforcement; *Tutorial Programs; Vocabulary Development
IDENTIFIERS *Project Follow Through

ABSTRACT

This study provided a review of literature related to peer tutoring and investigated the effects of three variables on peer tutor and peer tutee performance. The variables were (1) the achievement level of the tutor, (2) brief tutor training in reinforcement and corrective feedback procedures, and (3) tutor expectancy about tutee performance. A total of 112 high and low achieving second graders in a Follow Through program were selected as tutors. Half of these tutors were randomly selected to receive two half-hour training sessions. The training was conducted in a small group setting and emphasized two reinforcement and feedback procedures: (1) responding to a correct answer with a positive comment, and (2) responding to an incorrect answer or no answer by providing the correct answer and giving the tutee a chance to say the correct answer. A group of 112 first graders was randomly selected to serve as tutees. Instructions to tutors prior to the tutoring session contained either high or low expectancies about the tutees' academic performance. Both tutors and tutees were pre- and posttested on 30 sight word flash cards, with which the children worked during the tutoring session. Results showed that both tutors and tutees learned a significant number of words. Tutors' achievement level was not a significant factor, as measured by words learned, but tutees tutored by trained tutors performed better than tutees tutored by untrained tutors. Tutor expectancies did not affect tutee performance on the posttest. (Author/BRT)

ARIZONA CENTER FOR EDUCATIONAL
RESEARCH AND DEVELOPMENT
College of Education
University of Arizona
Tucson, Arizona 85721

F. Robert Paulsen, Dean
College of Education

Marsden B. Stokes, Director
Arizona Center for Educational
Research and Development

Joseph M. Tillerup, Director
TEEM Follow Through Project

THE EFFECTS OF TUTOR ACHIEVEMENT LEVEL, REINFORCEMENT TRAINING, AND EXPECTANCY ON PEER TUTORING

by: Eva E. Conrad

The research reported herein was supported by funds from the Follow Through Division, U. S. Office of Education.

July 18, 1975

ACKNOWLEDGMENTS

The author wishes to express appreciation to Dr. Joseph M. Fillerup and to Dr. Robert K. Rentfrow, Arizona Center for Educational Research and Development, for their moral and financial support, and to Dr. Ronald Henderson, supervisor of this study, Dr. Keith Meredith, Dr. John Bergan, Dr. Ralph Wetzel, and Dr. Wayne Carroll for their advice and encouragement. Thanks are also extended to the administrators, teachers, and first and second graders in Lincoln, Nebraska. The patience and sacrifice of Rex and Lettie Conrad made this research possible.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

ABSTRACT

Open classroom systems frequently use peer tutoring techniques as a means of individualizing instruction. Peer tutoring provides an opportunity for a one-to-one relationship within an academic context and is often a spontaneous outcome of heterogenous groupings. However, research-derived guidelines for optimal peer tutoring efficacy are lacking. This study investigated the effects of three variables on tutor and tutee performance: (1) the achievement level of the tutor, (2) brief tutor training in reinforcement and corrective feedback procedures, and (3) tutor expectancy about tutee performance.

One hundred and twelve high and low achieving second graders in a Follow Through program were selected as tutors. Half of these tutors were randomly selected to receive two half-hour training sessions. The training was conducted in a small group setting and emphasized two reinforcement and feedback procedures: (1) to respond to a correct answer with a positive comment, and (2) to respond to an incorrect or no answer by providing the correct answer and giving the tutee a chance to say the correct answer. One hundred and twelve first graders were randomly selected to serve as tutees. Instructions to tutors prior to the tutoring session contained either high or low expectancies about the tutees' academic performance. Both tutors and tutees were pre- and posttested on 30 sight word flash cards, with which the children worked during the tutoring session. A behavioral observation instrument

provided a measurement of tutor teaching behavior: number of cards presented, type and frequency of corrective feedback, positive reinforcement, negative gestures or comments, and providing or accepting a word incorrectly.

Analysis of variance procedures were used to analyze data provided by the two dependent variables. Both tutors and tutees learned a significant number of words. Although this study does not compare peer tutoring with other instructional techniques, benefit to both tutors and tutees from peer tutoring is demonstrated in the data. Tutors' achievement level was not a significant variable as measured by tutees' pre- to posttest improvement. Although there was virtually no difference between the two groups of tutees on the pretest, tutees tutored by trained tutors performed significantly better on the posttest than did tutees tutored by untrained tutors. These data illustrate that an increase in tutoring skills after a minimum of tutor training results in increased achievement. Other indications of the effectiveness of the brief tutor training are that trained tutors demonstrated a significantly higher frequency of the behavioral measures of corrective feedback, and that they provided significantly more verbal reinforcement than did untrained tutors. Training of tutors is obviously a crucial issue for effective peer tutoring. There was a significant difference between the pretest scores of tutees about whom tutors had low expectancies and tutees about whom tutors had high expectancies. However, this difference is not evident in posttest data. The fact that posttest scores

did not reflect the differences seen in the pretest can be interpreted to indicate that tutors were unbiased in their tutoring behavior and uninfluenced by the experimentally-manipulated expectancies about tutees' performance.

Peer tutoring guidelines that can be generated from this research include: (1) all children in a class, regardless of achievement level, should be selected to serve in the tutoring role; (2) brief tutor training in basic reinforcement and corrective feedback procedures is essential to an effective peer tutoring program; and (3) expectancies about tutees' performance may result in less biased teaching behavior by peer tutors than by adult tutors.

CHAPTER 1

INTRODUCTION AND BACKGROUND RESEARCH

Tutoring is an educational tool dating from the time of Plato. Rekindled interest in tutoring is largely a result of compensatory education programs and open classroom innovations.

A commonality of the divergent compensatory education programs developed in the past decade is an increase in the quantity of education provided to disadvantaged children (Tannenbaum, 1968). Preschool programs, longer school years, smaller teacher-child ratios, more teaching aids, and open space schools contribute to a goal of individualizing education. These innovations result in a need for increased manpower in the classroom. This need is often unfulfilled due to a paucity of qualified personnel and/or funds. Tutoring by peers or paraprofessionals has proven in many cases to be a viable solution (e.g., Mollod; 1970).

Emphasis on the individual is a salient characteristic of open education (Barth, 1972). Individualized instruction permits children to learn by discovery and at their own pace. Wolfson (1972) believes that individualization can lead to less stress on conformity and to a promotion of initiative and creativity. The resultant shifting away from large group teaching and from group norms requires lower teacher-pupil ratios. Paraprofessional or peer tutoring frees the teacher to concentrate his or her skills on individualization of instruction.

Heterogenous grouping of children is another distinctive feature of open classrooms. This procedure of grouping children of different skill levels or ages optimizes the opportunity for spontaneous or structured peer tutoring.

Peer tutoring may be a vehicle for teaching "generative" skills to tutors (Resnick, 1972). A generative skill is one that can facilitate new learning. For example, if a tutor is required by the tutoring process to organize materials inductively to enhance the tutee's acquisition, then the generative skill of inductive organization is available to the tutor in his future learning. Such skills are subsumed under the rubric of learning how to learn and provide students with generalized skills for processing new information and in using acquired skills.

Adults as Tutors

Education is following the lead of medical and social work professions in the use of paraprofessional and volunteer personnel. Paraprofessionals in the classroom lessen the teacher's workload and free teachers' time, which can then be devoted to individualizing instruction. Each child in the classroom can receive increased individual attention as the number of teachers or tutors increases.

Tannenbaum (1968) reports that parents compensate for their lack of professional training with persistence, dedication, and frequency of contact with the children. Supplementary Teaching Assistance in Reading is a program which uses both paraprofessionals and parents to raise the literacy level of underprivileged children.

Paraprofessionals indigenous to a largely Spanish-speaking neighborhood were trained in three reading readiness activities: code breaking, formal language experience, and visual perceptual experiences. Twenty-eight aides taught these skills to the volunteer parents in their own homes. Parents were provided with materials for lessons and with structured guidelines for fifteen weeks. Compared to controls and to a group of children who attended weekly remedial sessions with a teacher, the children participating in the home tutoring sessions achieved the highest mean scores on the Frostig Developmental Tests of Visual Perception and the Metropolitan Reading Readiness Tests. The author suggests that the lack of significance was a result of unequal group N's and the inability to test all subjects due to funding cutbacks. In another study using trained volunteers (Murray, 1972) parents taught their second graders at home for twelve one-hour sessions. Standardized reading test scores were significantly higher ($p < .0005$) for the tutored group compared to untutored controls.

Ellson, Barber, Engle, and Kampworth (1965) provided adult tutors with programmed materials in reading. An alternation of tutoring and classroom instruction resulted in significantly superior achievement compared to classroom instruction alone. The trend for tutoring alone as superior to classroom instruction alone was not statistically significant. Although not evident in all cases, there was an ability level interaction with the poorest readers realizing superior gains in the tutor condition.

Shaver and Nuhn (1971) identified forty-six underachievers in grades four, seven, and ten via the Sequential Tests of Educational

Process. These children were tutored by volunteer adults in reading and writing for one hour per day in either a one-to-one or a one-to-three adult/child ratio. Tutored subjects made significantly greater gains on the California Tests of Mental Maturity than did comparably underachieving controls. In a two year follow-up, the mean gains were sustained for the older students but not for the fourth graders. In the only significant gain score-group size interaction, tenth graders realized greater gains in reading when tutored in a one-to-one, rather than in a one-to-three adult/pupil ratio.

Bausell, Moody, and Walzl (1972) investigated the effects of tutoring by trained teachers versus untrained undergraduates. The tutors were supplied with eight mathematics objectives, but the instructional procedures were unspecified. Controlling for the total amount of instruction time, control subjects received regular classroom instruction in mathematics. The 120 tutored fourth and fifth graders made significantly ($p < .05$) greater gains than did the controls. Teacher training and tutor ability level were not significant variables.

Hamblin and Hamblin (1972) studied the effects of adult versus peer tutoring and contingent versus noncontingent reinforcement in a program designed to teach reading to disadvantaged preschoolers. The thirty-two children were divided into four groups: peer tutoring plus contingent reinforcement, peer tutoring plus noncontingent reinforcement, adult tutoring plus contingent reinforcement, and adult tutoring plus noncontingent reinforcement. The dependent variables were the number of books read, the number of words learned, and the number of symbols learned. High to medium IQ preschoolers were superior in all

three measures under peer tutoring plus contingent reinforcement, with peer reinforcement plus noncontingent reinforcement resulting in the second highest scores of two of the three dependent variables. Low IQ preschoolers realized the greatest gains under peer tutoring, with the contingency of reinforcement creating only negligible differences. The authors also report significant gains in reading for the peer tutors.

Peers as Tutors

Peer tutors provide the same benefit of increasing teacher time for individualizing instruction as do adult paraprofessionals and volunteers. In addition, peer tutoring creates an opportunity for the utilization of newly acquired skills in a meaningful context. The helper therapy principle embodies the notion that a helper benefits from the helper role as well as from the content of the situation (Durlack, 1973).

Tanner and Lindgren (1971) conclude from an extensive cross-age peer tutoring program that when one teaches, two learn. These authors suggest that a true test of a child's understanding of a precept or principle is his ability to teach it to another child. Feldman and Allen (1972) demonstrated the cogent effects of the peer tutor role. Low achieving fifth graders tutored third graders in mathematics or spent equal time alone. The peer tutors learned significantly more than did their classmates who studied alone.

Von Harrison, Nelson, and Tregaskis (1972) compare structured peer tutoring to programmed instruction. However, in the areas of cost and flexibility, structured peer tutoring is deemed by the authors a

superior form of individualized instruction. Reasons cited for this superiority include ability to monitor tutee's oral response and tutee's interaction with the instructional materials. Analogous to other instructional options, peer tutoring is a technique adaptable for use in a variety of academic content areas.

Peer tutoring typically involves the use of older-younger dyads or same-age dyads. Variables which have been investigated in peer tutoring include tutor training, nature of the task, and extrinsic reinforcement.

Unstructured programs and programs with no tutor training have not generally met with success. Hunt (1971) attempted to assess the effects of tutoring on the racial attitudes and academic achievement of high school sophomores. Students were permitted to volunteer to serve as a tutor or tutee twice a week in an academic topic of their choice. Neither the experimental nor the control group evidenced significant achievement gains. As an explanation for the lack of significant achievement gains the author cites the problem of maintaining quality control on the tutoring activities and the paucity of valid short-term achievement and attitude assessment instruments. Recommendations for future programs include a precise definition of the role of the tutor and concrete educational goals for the tutoring session.

Greenwood (1973) assigned elementary school children with histories of management problems to groups of five on the basis of Sullivan Mathematics Placement Scores. The tutors for each group were selected by balloting of the group. The dependent variables were the frequency of inappropriate tutor behavior and the number of incorrect

written workbook responses. An ABA design was comprised of baseline, tutor training, and return to baseline. Tutor training included instructions and modeling of contingent, positive reinforcement. There was a significant difference in the positive reinforcement. There was a significant difference in the positiveness and appropriateness of the peer tutors' reinforcement, but no significant achievement gains were reported. Tutor training did not include instruction in the content area of mathematics nor specific mathematics objectives.

Lippitt, Eiseman, and Lippitt (1969) conducted a cross-age peer tutoring program. Sixth graders received training, via role-playing and discussion, in the constructive versus destructive use of errors. An error is used constructively if the source of the error is analyzed and information provided on the tutee's academic needs is used in further instruction. This is opposed to a destructive, less productive right-or-wrong approach. After this initial training, which also emphasized the tutor-tutee relationship as a partnership, the older pupils decided mutually with the first to fourth grade teachers the content of the tutoring sessions. Results of the program were measured by both achievement and self-concept tests. Both tutors and tutees showed positive increases in academic learning, self-concept, social acceptability, self-discipline, and attendance than did matched controls.

A second cross-age peer tutoring program (Hagen and Moeller, 1971) followed a similar peer tutoring training procedure. Both programs stress the importance of using tutors at least three years older than the tutee. Training of tutors is typically conducted in a seminar

with tutees' classroom teachers, and issues discussed are appropriate materials and the handling of discipline problems. The purpose of this research was to investigate the teaching process and the social interaction in the tutor-tutee relationship. The authors conclude that cross-age peer tutoring is a viable teaching option and that both participants benefit socially from peer tutoring.

Niedermyer and Ellis (1970) developed a peer tutor training program which includes training on correctly confirming tutee responses, praising the tutee, showing the tutee the correct response if an error is made, eliciting a correct response after an error is made, rephrasing questions if no response, and avoiding negative verbal behavior. Instructions, modeling, and role-playing were used in tutor training. During tutoring, clearly structured guidelines were followed by the tutors. Recommendations from this pilot training program stress specific behavioral educational objectives and an easily monitored system for teacher appraisal.

Structured programs have been more successful in realizing tutor and tutee achievement gains. Von Harrison et al. (1972) view structured tutoring as a teaching technique which can be adapted to any academic content area. In their supplementary tutorial reading program, supervisors trained and monitored tutors. Tutors were trained in positive verbal reinforcement and feedback as well as in the specific content procedures of naming the letters of the alphabet, production of the common consonants, vowels, digraphs, and the phonetic decoding of nonsense words. The tutors ranged from low to high reading achievement and were in grades four through six. The authors' criteria of 80-80

(80% of the subjects attaining 80% or above on the posttest) was not met. From the total sample of non-reading second grade tutees, 72% scored 80% or above on the posttest, 20% scored between 60% and 80%, and 8% scored between 0 and 60%.

A rather unique peer tutoring program (Amster, Davis, and Logan, 1975) trains peer tutors in basic mathematics skills and in abstract curriculum. The ninth grade tutors then teach mathematics to whole classes of seventh graders twice a week. This technique is compared to three other teaching options: (1) classes taught twice a week by an adult math specialist, (2) standard curriculum taught by a teacher from another class twice a week, and (3) regular classroom teacher teaching standard curriculum five days a week. The peer tutoring group showed no special advantage over the other groups for tutors or tutees on tests of math gains, spatial reasoning, or Ravens Progressive Matrices. Female tutors did reveal increased gains in positive self-attitude compared to other ninth graders who had not engaged in tutoring.

Morgan and Toy (1970) instructed older elementary school children to create a "warm, friendly, and accepting atmosphere" in the tutoring sessions. Details of the content to be covered with the tutees were determined by the classroom teachers. Academic achievement in arithmetic, spelling, and reading was assessed for both tutors and tutees by the Wide Range Achievement Test. Both experimental and control tutees made significant ($p < .05$) academic gains. Classroom teachers in this traditional school spent the time freed by the tutors in individualizing instruction for the control subjects. The most striking differences were in the achievement gains made by the tutors:

$p < .01$ for reading and spelling, and $p < .05$ for arithmetic. Experimental-control children's achievement gains were equally high in another study (Stainback, 1972) due to an increment in teacher attention to the control children.

The nature of the task in a tutorial session is another determinant of success. An additional recommendation provided by Niedermyer and Ellis (1970, p. 28) is the use of "stimulus-response learning tasks that require practice and repetition." Advanced second and third graders tutored classmates in both word knowledge and comprehension skills for four months (Mollod, 1970). With the tutors improving equally well as the tutees, the pre- to posttest mean gain scores were significant ($p < .005$) for word knowledge, but the mean gain scores in comprehension were not significant.

Evidence for positive effects of extrinsic reinforcement in tutoring situations is scant. Forty disadvantaged first through third graders were tutored on a one-to-one basis by fifth and sixth graders (Snapp, Oakland, & Williams, 1972). After tutor training, the dyads were randomly divided into two groups: one received reinforcement while being tutored in reading, and one received no reinforcement. Tutoring resulted in reading gains, but the pre- to posttest differences were not significant. The reinforcement group was not superior to the non-reinforcement group. Hamblin and Hamblin (1972) found that with high to medium IQ children, tokens had only a slight effect on book reading and word and symbol acquisition. Peer or adult tutoring plus contingent reinforcement was more effective than peer or adult tutoring plus non-contigent reinforcement. However, peer tutoring was more effective with

low IQ subjects regardless of the reinforcement contingencies. Brown (1972) had adolescent tutors dispense no reinforcement, contingent reinforcement, or noncontingent reinforcement while tutoring younger children in reading. Both reinforcement groups scored slightly higher than the no reinforcement group on a 100-word recognition test although the differences were not significant.

Myers, Travers, and Sanford (1965) provide experimental support for the importance of interaction in tutoring to achieve mutual learning. Fourth through sixth graders were randomly divided into one of four groups to learn German words: tutor provides feedback to tutee; tutor-tutee mutually interact with stimulus materials; tutor-tutee switch roles halfway through the study; child works alone with stimulus materials. In both immediate and delayed tests, the first group produced the least effective learning. Differences among the other three groups were not significant, although anecdotal reports indicate motivation was highest in the third group. Contrary to other studies, in this experiment the tutor brought no particular process or content skills to the tutorial situation.

The literature review did not locate a study in which tutor achievement level was manipulated as an independent variable. Perhaps in keeping with a "common sense" approach, studies of same-age peer tutoring employ the higher achieving peer in the tutor role. Hamblin and Hamblin (1972) report significant reading gains for both tutors and tutees when higher achieving preschoolers tutored their classmates. The third grade children made greater gains when paired with a partner close to them in achievement level, although this trend was not

statistically significant. Guarnaccia (1973) employed higher achieving third and fourth graders as tutors for their classmates. The peer tutoring sessions in arithmetic resulted in a mean increase of six months on achievement tests for both tutors and tutees compared to control classrooms. This difference was maintained in a two month follow-up. Although tutor achievement level was not an independent variable in this study, the author notes that regardless of the tutee's achievement level, greater gains were realized the higher the tutor's initial achievement level. This is contrary to Cloward's (1967) conclusion in a review article that the intellectual credentials of the tutor are not positively correlated with his effectiveness.

Benefits to Tutees and Tutors

An optimal learning situation has been defined (Hamblin, Buckholdt, Ferritor, Kazloff, and Blackwell, 1971) to include reinforcement for the adaptive behavior, rapid and consistent feedback, and self-pacing wherein the required response neither slows nor rushes the learner. Structured peer tutoring can create such an optimal situation for the tutee. Instructional materials ideally are geared to the tutee's specific academic needs. The tutee has an opportunity for immediate corrective feedback and for the observation of a peer model's interaction with academic material within a situation which carries no stigma for giving an incorrect answer.

Attitudinal as well as academic benefits can be derived for the tutee. Evaluative feedback and specific criteria can be particularly beneficial to children low in willingness to attempt difficult tasks.

Children were most willing to attempt a difficult motor task when working in peer pairs than when alone or in the presence of the entire class (Torrance, 1969). Von Harrison et al. (1972) report parental anecdotes of attempts to read and sound out words at home after a peer tutoring program for non-readers. Two studies provide evidence of lower IQ children (Hamblin and Hamblin, 1972) and low achievers (Ellson et al., 1965) realizing greater gains than classmates from tutoring as assessed by criterion-referenced measurement techniques (e.g., number of words read, number of books read, etc.)

Especially beneficial in communities where teacher-pupil rapport is difficult to establish, the older school-age child can provide a model for sociological as well as cognitive development (Bernard, 1972). Older peer tutors share the same student status at their younger tutees, often yielding respect without an association of past punishment (Morgan and Toy, 1970).

For the tutor, peer tutoring is an opportunity to consolidate learning and to develop and improve communication skills. The majority of the studies previously cited describe achievement gains that exceed or at least match tutee achievement gains. Cloward (1967) concludes that tutors should be carefully selected as they are the primary beneficiaries of the tutoring experience. Feldman and Allen (1972) used peer tutoring primarily to increase the academic skills of the tutors. Durlack (1973) reports a peer tutoring program for ninth graders who were experiencing from minor to severe school adjustment. These students worked in lower grade classrooms a full morning or afternoon five days a week. Although gains were not reported, the rationale of the

program was that peer tutoring provided the tutors with an opportunity to develop social competency, to adjust to working with both younger and older individuals, and to assume personal responsibilities.

Directions for Peer Tutoring Research

The goal of significance at $p < .05$ is not consistently realized in tutoring studies. The lack of control over total instruction time (Bausell et al., 1972), a small N (Ellson et al., 1965), inadequate evaluative procedures (Morgan and Toy, 1970), and the brief duration of many intervention programs contribute to this lack of significance. In all studies reviewed, however, there was a trend for tutoring to be superior to classroom instruction alone. When viewed within the context of the difficulties inherent in applied educational research (Bruner, 1965), this consistency across divergent settings and methodologies justifies further investigation of the variables in tutoring procedures.

Two tutoring variables which have been investigated but which do not yield readily applicable, conclusive results are those of tutor training and the effects of tutor achievement level.

In this review, research in which tutor training served as an independent variable compared an absence of training to training (Greenwood, 1973), or provided pilot test results of highly structured tutor training programs (Lippitt and Lohman, 1965; Von Harrison et al., 1972). Programs which provide no training to tutors in either technique or content have not been successful (Hunt, 1971). Greenwood (1973) compared the effectiveness of tutors after training to their

pre-training management behavior. Improvement in reinforcement procedures was significant. However, the tutors did not receive training in the content area of arithmetic, and significant improvements were not realized in tutee academic performance. Highly structured tutor training programs (e.g., Niedermyer and Ellis, 1970; Von Harrison et al., 1972) include instruction in both content area and in reinforcement techniques. At the conclusion of these time-consuming, costly tutor training programs, neither study was able to report significant gains for the tutees. A literature search revealed no studies in which the amount of tutor training in reinforcement techniques and/or content areas was manipulated as an independent variable.

This literature review failed to locate a study in which tutor achievement level was manipulated as an independent variable. Generally, in same-age peer tutoring, the achievement level of the child is the basis for the selection of the tutor (e.g., Hamblin and Hamblin, 1972; Mollod, 1970; Guarnaccia, 1973). In these studies, the selection of higher achieving students as tutors eliminated the need for tutor training in the content area. Significant achievement gains for the peer tutors as well as for the tutees have been reported in the above mentioned studies with both same-age tutors and with older-younger peer dyads (Morgan and Toy, 1970). Tutoring provides an opportunity for tutors to consolidate learning. Academic improvement by tutors may only be true for children who are high achievers prior to tutoring. Further research is needed to clarify the effects of middle or low achieving tutors on both tutee and tutor performance.

Expectancy

Teachers expectations have been defined (Brophy and Good, 1974) as inferences that teachers make about the present and future achievement and behavior of their students.

The evidence of teachers' expectancies of pupil performance affecting IQ scores presented by Rosenthal and Jacobson (1966) has led to further research and to discussions of teacher expectancies in teacher education programs. Rosenthal and Jacobson (1966) randomly selected 20% of an elementary school population in grades one through six. Teachers were advised that the Tests of General Ability revealed that these selected children would show unusual intellectual gains during the academic year. Intelligence testing at the end of the school year did show significant ($p < .02$) gains for the experimental subjects. There was a grade-gain score interaction, with the greatest effect present in the lower grades. Teachers rated the designated bloomers as more likely to succeed in the future and more intellectually curious than classmates. This effect was not significant and also was most pronounced in the early grades.

This study created a minor whirlwind of publicity and controversy. One of the more entertaining debates was the resultant one between Rosenthal and Thorndike. Thorndike (1968, p. 709) summarized his review of the original Pygmalion research with "Alas, it is so defective technically that one can only regret that it ever got beyond the eyes of the original investigators." Snow (1969) compiled a synthesis of the various complaints about the Rosenthal and Jacobson (1966) research:

1. There was not enough detail to properly critique the study, let alone replicate it.
2. The standardized tests used (Tests of General Abilities) were not normed for the youngest subjects.
3. Teachers administered the tests, therefore there was no standardized administration.
4. The dependent variable used in computation was gain scores, even though several mean pretest differences between treatment groups equaled or exceeded obtained posttest differences.
5. Post-experimental interviews indicated that teachers could not remember the names of those children designated at the beginning of the school year as "bloomers".

Rosenthal (Evans and Rosenthal, 1969) answered some of his critics by providing raw data for closer inspection and re-analysis. He states that the effects of teacher expectations on one variable (e.g., reasoning) are not dependent on the inclusion of the first and second grade data. The artificially low IQ (e.g., an IQ of 58 for first graders) was created by children omitting answers instead of guessing.

Replications of the original Pygmalion have yielded inconclusive results. Claiborn's (1969) study was very like that of Rosenthal and Jacobson (1966) with two exceptions. The bias was introduced in the second semester, after teachers had had an opportunity to form their own expectations. The second exception was that there was only a two month lapse between the pre- and posttesting on the Tests of General Ability. Claiborn's results were negative: neither process

nor product measures yielded any effect from the induced expectations. A second very close approximation of the original research also yielded negative results (Evans and Rosenthal, 1969). Differences between this study and the original were that it was conducted in two schools instead of one, and that the students were middle class, not lower class. Fine (1972) manipulated teachers' expectancies of 180 second graders and measured the effects of that expectancy on four dependent variables: teacher rankings of children, vocabulary test, reading comprehension tests, and IQ scores. Positive expectancy resulted in significantly higher reading comprehension scores and teacher rankings. Differences in vocabulary test scores and in IQ's were not significant. Fleming and Anttonen (1971) studied the effects of expectancy on the standardized test performance, teacher-assigned grades, and self-concept of 1087 second graders. Although the self-fulfilling prophecy was not supported, there was an SES-expectancy interaction, with the high opinion teachers producing greater academic gains in middle class children than did low opinion teachers.

In the preceding studies the teachers interacted with their pupils for entire school terms armed with the biased expectancies. Only minor support of positive teacher expectancies resulting in positive academic gains can be found in these studies. The effects of manipulation of teacher expectancy are even more evident in studies of shorter duration. One could conclude that teachers rely on presented "factual" material about their pupils to a greater extent when the opportunity for long-term, personal observation is lacking. Beez (1970) gave biased high ability or low ability psychological reports

to 60 teachers prior to a one-to-one, ten minute tutoring session. After working with the Head Start children for ten minutes with reading flash cards, the children were posttested, and the teachers were asked to rate the child's achievement level, social competency, and intellectual ability. During the tutoring session, various behavioral observations were made, such as the number of cards presented, the amount of time spent on each card, and the number of times a word was read to the child. Teachers' behavior was significantly affected by the biased psychological reports. Teachers expecting high ability presented significantly more words to their tutees than did the other teachers. Obviously, the number of words learned by the two groups of tutees was significantly different. On all three measures (achievement level, social competency, and intellectual ability) the high expectancy teachers rated their children significantly higher than did low expectancy teachers. Low expectancy teachers explained the words significantly more often than did high expectancy teachers. These data are consistent with other research (Brophy and Good, 1974) which indicate that teachers demand better performance from those students for whom they hold high expectations.

Pellegrini and Hicks (1972) randomly selected elementary school children to be labeled as high (120-129), average (95-105), or low (85-95) IQ for college volunteer tutors. One group of tutors who worked with a second group of "high" IQ children were familiar with the tests with which the children would be re-evaluated at the end of the seventeen week program. "High" IQ children gained more than the "average" or "low" groups, but the difference was not statistically

significant. The second group of "high" IQ children scored significantly higher on the association test (Peabody Picture Vocabulary Test). However, this group was not significantly different on the more conceptually difficult task (similarities subtest on the WISC).

Mason (1973) used biased psychological reports to test if expectancy affected the observations and/or expectations of 79 teachers. After receiving a negative, neutral, or positive psychological report, the teachers watched a video-tape of a male or female kindergarten child taking a test of concept development. The teachers were asked to rate the child as passing or failing, based on their observations, and to predict the child's grades at the end of the first grade. The biased reports did significantly affect teachers' predictions but did not affect their observations.

Results of research and conclusions drawn by research reviewers are less than conclusive. Grieger (1971, p. 90) reviewed the literature on teacher expectancy and concluded that it is "more difficult to demonstrate and less pervasive than has been claimed." However, Brophy and Good (1974) conclude that regardless of one's views on the original Rosenthal and Jacobson (1966) research, there has been sufficient research with a variety of situations and subjects to conclude that teachers' expectations can and do function as self-fulfilling prophecies, although the process is far from automatic.

Research which provides process measures yields information on the relevant variables in the communication of teachers' expectations to students. Although Conn, Edwards, Rosenthal, and Crowne (1968) failed to replicate Rosenthal and Jacobson (1966), they did find that children

who could discriminate emotional components in an adult female's voice showed a significantly greater advantage from positive teacher expectations than did other children in the experimental group. Anderson and Rosenthal (1968) randomly designated a group of retarded children as late bloomers to counselors in a summer camp. The dependent variables were a general abilities test and a self-help measure. Process data indicates that camp counselors devoted less time to doing unnecessary tasks for the "bloomers", allowing them to practice doing things for themselves. Rubovits and Maehr (1971) found that undergraduate volunteer teachers working in a microteaching situation initiated more interactions with, requested more statements from, and praised more the children randomly designated as gifted, as opposed to the children labeled normal. Medinnus and Unruh (1971) selected two boys in normal IQ range (95-105) from twenty Head Start classrooms. Teachers were told that one boy was high-ability, with an IQ above 105, and that the other boy was low-ability, with an IQ below 105. Teachers then worked individually with students from their own class on a block-sorting task. The dependent variables were frequency of contact, praise, cooperation, criticism, direction, and interference from the teacher during the individual work session. There were significant differences in teacher behavior from the high-ability to the low-ability tutoring sessions on only two variables. While tutoring children labeled as high-ability, teachers provided more praise and less criticism than during sessions with the designated low-ability children.

The Medinnus and Unruh (1971) study is unique in that significant process differences can be attributed to experimentally-induced

expectancies even though the teacher had enough previous experience with the children to form his own expectations. Credibility of the experimentally-induced expectations may be a weakness of many studies in this area. The experimental "factual" information may clash with the child's past or present performance and/or with the teacher's expectations of how a particular child or type of child will perform.

Brophy and Good (1974) summarize the process variables that have been found to correlate with the communication of positive expectations for student performance: (1) frequency of contact, (2) positive affect, (3) difficulty of material presented, (4) attentiveness, and (5) persistence in seeking a response. Expectancy may also be a partial determinant of teaching behavior in peer tutoring situations. The literature search revealed no studies in which expectancy was manipulated in peer tutoring.

Summary

Several practically relevant questions remain unanswered in the area of peer tutoring. This literature review revealed no studies which assessed the effects of tutor achievement level on both tutor and tutee performance. Past research does indicate that structured programs which include tutor training are the most successful. Two issues, however, are unresolved: the minimum amount of training that can yield positive tutee and tutor gains, and the relative importance in tutor training of process and content skills. A third unanswered issue is the potential effect of tutor expectancies about the tutee on tutor teaching behavior and, consequently, on tutee gains from peer tutoring.

This study investigates the effects on peer tutoring of three variables: (1) tutor achievement level, (2) a small amount of tutor training in process skills, and (3) experimentally-induced expectancies.

CHAPTER 2

GENERAL PURPOSES AND HYPOTHESES

The purpose of this research was to investigate the effects of tutor achievement level, reinforcement training, and expectancy of tutee performance on tutors' and tutees' gain scores and on tutors' teaching behavior.

The predicted direction of effects stated in the following hypotheses reflect the viewpoint of the author.

Hypothesis 1. Tutee pre-to posttest gain scores would not be significantly different as a function of tutor achievement level.

Hypothesis 2. Tutor achievement level would result in significantly different tutor gain scores, with low achievement tutors making greater gains than high achievement tutors.

Hypothesis 3. Tutor reinforcement training would result in significantly higher tutee gain scores.

Hypothesis 4. Tutor reinforcement training would not result in significantly higher tutor gain scores.

Hypothesis 5. Expectancy of tutee performance by tutors would not result in significantly different tutee gain scores.

Hypothesis 6. Expectancy of tutee performance by tutors would not result in significantly different tutor teaching behavior (e.g., number of cards presented, etc.).

Hypothesis 7. There would be a significant relationship between: (1) the number of words learned by the tutees and the frequency of positive verbal reinforcement, and (2) the number of words learned by tutees and the number of cards presented.

Hypothesis 8. Tutors who received reinforcement training would emit the following two behaviors with a significantly higher frequency than would the tutors who receive no training: (1) positive verbal reinforcement, and (2) the provision of a correct word or picture if an incorrect answer or no answer is given.

Hypothesis 9. There would be no significant sex effect among the treatment groups.

CHAPTER 3

METHOD

Subjects

The subjects were 224 first and second graders attending Lincoln public schools. All children were currently in Follow Through classrooms which implement the Tucson Early Education Model.

Procedure

The second graders served as tutors. Scores on the Metropolitan Achievement Test administered at the end of the 1974 school year were divided into high and low groups. From the available sample of 198 scores, the 28 highest and 28 lowest males and females (N = 112) who were still in the school district were identified.

Sampling evenly from the population of high and low achievement males and females, the tutors were randomly selected to recieve reinforcement training or to receive no training. The author used discussion, modeling, and feedback during role-playing to instruct the tutors in two basic teaching procedures:

1. Respond to each correct answer with a positive statement (e.g., "good", "that's right").
2. Respond to each incorrect answer by providing the correct answer and repeating the question.

Within the same week, the tutors recieved two thirty-minute training sessions in groups of four. The training took place in a room separate from their classroom.

Within the populations of same sex and school as the tutors, the tutees were randomly selected from the 180 available first graders. Both tutors and tutees were pretested on 30 word cards randomly selected from a set of 118 Milton Bradley Picture Flash Word Cards for Beginners, Grades 1-3. Sampling evenly from the populations of high and low achievement levels, training or no training, and males and females, tutors were randomly divided to receive either low or high expectancy instructions, after pretesting and immediately before the peer tutoring session. Table 1 presents the number of children in each cell across the three independent variables.

Table 1. Distribution of Subjects Across the Three Independent Variables.

		Reinforcement Training		No Training	
		High Achiev.	Low Achiev.	High Achiev.	Low Achiev.
Low Expectancy Instructions	Males	Tr=7 Te=7	Tr=7 Te=7	Tr=7 Te=7	Tr=7 Te=7
	Females	Tr=7 Te=7	Tr=7 Te=7	Tr=7 Te=7	Tr=7 Te=7
High Expectancy Instructions	Males	Tr=7 Te=7	Tr=7 Te=7	Tr=7 Te=7	Tr=7 Te=7
	Females	Tr=7 Te=7	Tr=7 Te=7	Tr=7 Te=7	Tr=7 Te=7

Tr = Tutors
Te = Tutees

The low expectancy instructions were:

> Since you knew so many of these words, I am going to let you be the teacher. You will be the teacher and help a first grader learn some of these words. If you don't know a word, the picture on the back of the card will help you. You are very smart and know some of these words. But the first grader you are going to help doesn't know as many words as you. These words will be hard for him (her). But even though these words will be hard for him (her), try to help him (her) learn some of the words. Work with him (her) for ten minutes. If you go through all the cards, you can start over again, or you can just work on a few cards. It is up to you. You are the teacher. I will tell you when to stop. Remember, even though these words will be hard for him (her), try to help him (her) learn some of the words.

The high expectancy instructions were:

> Since you knew so many of these words, I am going to let you be the teacher. You will be the teacher and help a first grader learn some of these words. If you don't know a word, the picture on the back of the card will help you. You are very smart and know some of these words. The first grader you are going to teach also knows lots of words. These words will be easy for him (her). Since you both are smart, try to help him (her) learn some of these words. Work with him (her) for ten minutes. If you go through all the cards, you can start over again, or you can just work on a few cards. It is up to you. You are the teacher. I will tell you when to stop. Remember, even though these words will be easy for him (her), try to help him (her) learn some of the words.

While the author pretested and instructed the tutor, a second experimenter (also a female adult Anglo) pretested the tutee. After the pretesting and instructions, the tutor and tutee were brought together in a room and told to work on learning some new words. The flash cards used during the peer tutoring were the 30 cards used for pretesting and posttesting. During the tutoring, one of the two experimenters recorded tutor behavioral data (see Appendix A). A minimum interobserver percent agreement of .85% was achieved on all items on the observation form prior to the beginning of this research. Approximately 20% of the tutoring sessions were recorded by both observers. At the

end of ten minutes, the tutor and tutee played with a pegboard game for five minutes. The tutor and tutee were then posttested on all 30 flash cards.

CHAPTER 4

RESULTS

The number of sight words tutors and tutees got correct on the pretest and the posttest was analyzed by two five-way analyses of variance. The one within-subjects repeated measure was the score (pretest/posttest), and the four between-subjects variables were sex (male/female), achievement (high/low), expectancy (high/low), and training (yes/no).

An analysis of variance yielding a significant interaction requires the qualification of significant main effects (Kirk, 1968). If a graph of the interaction is ordinal, main effects are easier to interpret because the direction of change across the variables plotted on the abscissa is the same (e.g., Figure 1). However, a disordinal graph (e.g., Figure 2) necessitates qualified interpretation of significant main effects. A significant interaction indicates that one treatment behaves differently under different levels of the other treatment. Interpretation of an interaction requires post hoc analysis of simple effects, e.g., differences among cell means. Discussion of these cell means differences results in qualified main effects statements.

The source table for the analysis of variance on tutor scores are presented in Table 2. Training of tutors had a significant ($p < .01$)

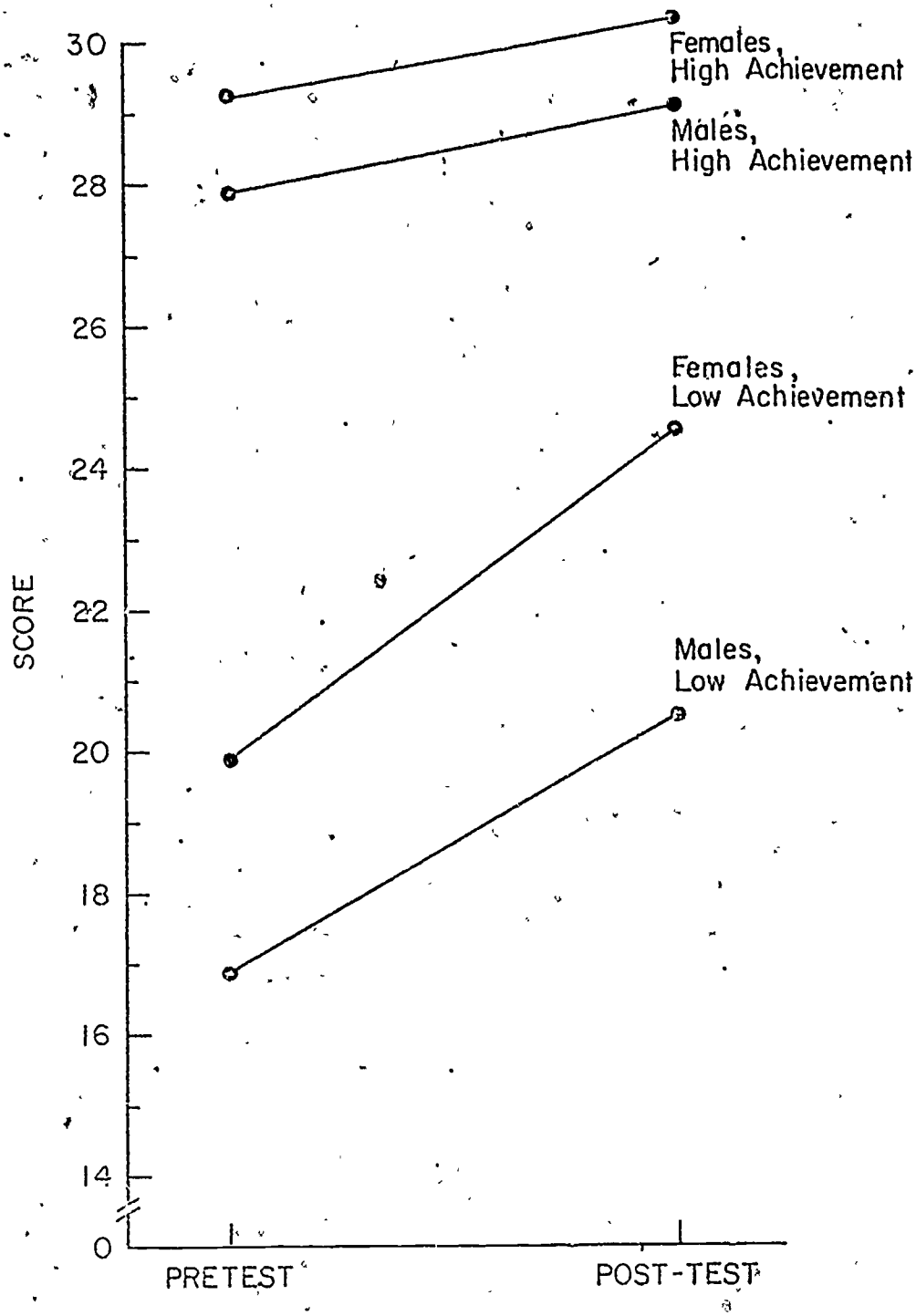


Fig. 1. Tutor Scores x Pretest/Posttest x Sex x Achievement

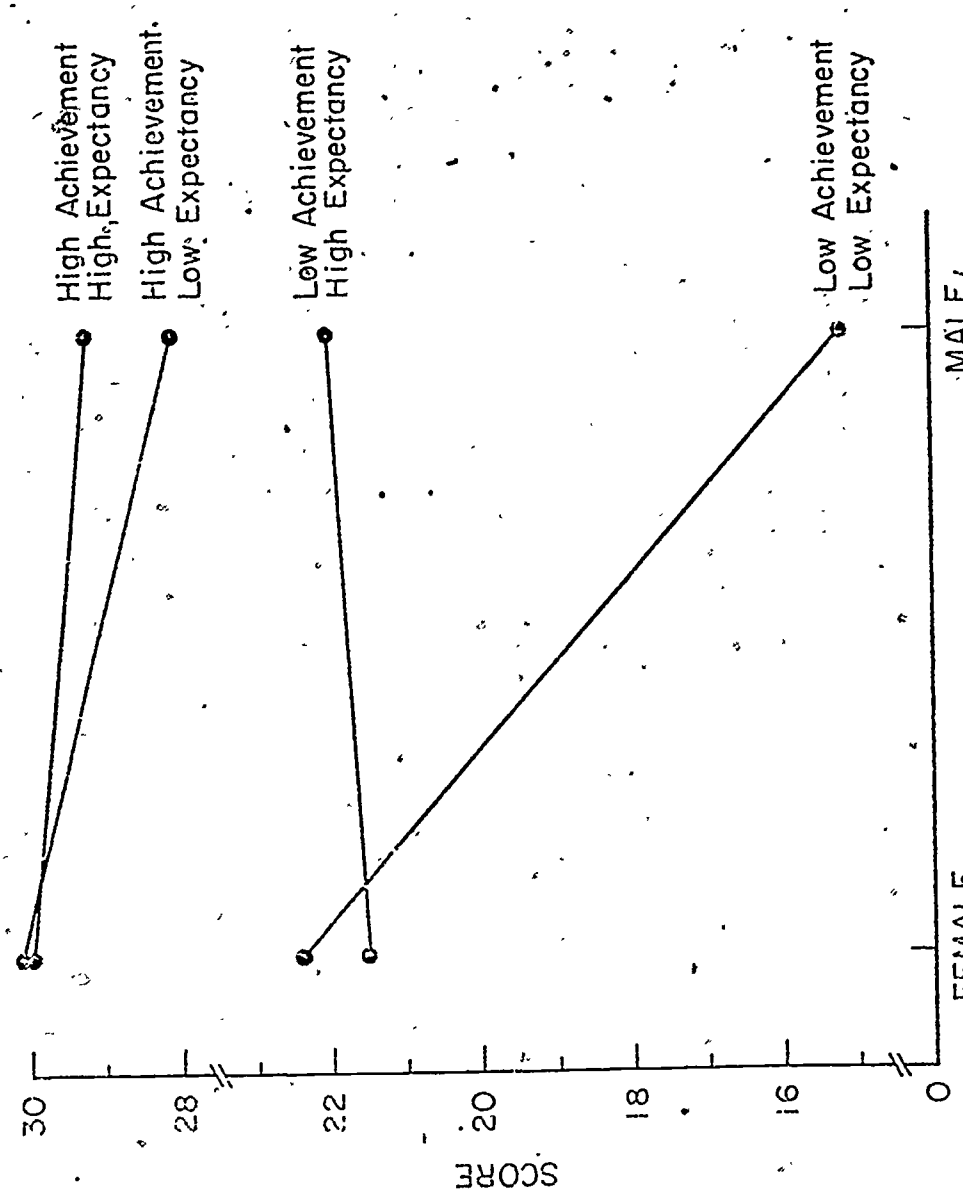


Fig. 2. Tutor Scores x Sex x Achievement x Expectancy

Table 2. Source Table for Analysis of Variance: Tutors

Source	DF	Mean Square	F
P	1	347.504	135.74**
F	1	290.290	6.67**
A	1	4174.504	95.90**
E	1	149.504	3.43
T	1	191.290	4.39**
PF	1	.361	<1.00
PA	1	149.504	58.40**
FA	1	81.361	1.87
PE	1	2.790	1.09
FE	1	281.254	6.46**
AE	1	81.361	1.86
PT	1	.004	<1.00
FT	1	124.504	2.86
AT	1	81.361	1.87
ET	1	4.290	<1.00
PFA	1	12.540	4.90**
PFE	1	.004	<1.00
PAE	1	.754	.30
FAE	1	206.361	4.74**
PFT	1	1.968	<1.00
PAT	1	3.754	1.47
FAT	1	107.254	2.46
PET	1	.754	<1.00
FET	1	.219	<1.00
AET	1	.754	<1.00
PFAE	1	.219	<1.00
PFAT	1	1.969	<1.00
PFET	1	4.861	1.90
PAET	1	.754	<1.00
FAET	1	.004	<1.00
PFAET	1	.754	<1.00
Within-Ss Error	96	2.563	
Between-Ss Error	96	43.527	

**$p < .01$

P = Pretest/Posttest
F = Sex
A = Achievement
E = Expectancy
T = Training

effect on tutors' scores on the association task. Tutors who had been trained had a higher mean score (25.7) than tutors who had not been trained (23.8).

Figure 1 graphically presents the ordinal interaction data for tutors' pretest/posttest scores by achievement and sex ($p < .01$). A post hoc analysis of these data (Table 3) provides evidence that the low achieving tutors were significantly lower in both pretest and posttest scores than high achieving tutors. The main effect for achievement was significant ($p < .01$). Also, low achieving tutors of both sexes made significant increases in their scores from pretest to posttest. The main effect for pretest/posttest and the pretest/posttest by achievement were also significant ($p < .01$). High achieving tutors demonstrated increases from pretest to posttest. However, due to a ceiling effect (maximum score = 30), these differences were not significant.

The sex by achievement by expectancy interaction in tutor scores ($p < .01$) as analyzed by Tukey's honestly significant differences test, (Kirk, 1968) is presented in Table 4. This interaction is illustrated in Figure 2. The post hoc test indicates that the only significant differences are within the high versus low achievement comparisons. Both the sex and expectancy variables failed to reveal consistent patterns.

The number of sight words correctly read by tutees before and after the peer tutoring session was analyzed by an analysis of variance. These results are presented in Table 5.

Table 3. Differences Among Tutor Means: Pretest/Posttest x Sex x Achievement

	16.9	19.8	20.5	24.5	27.9	29.1	29.3	29.8
Low Achievement								
M-Pretest 16.9		2.9*	3.6*	7.6*	11.0*	12.2*	12.4*	12.9*
F-Pretest 19.8			.7	4.7*	8.1*	9.3*	9.5*	10.0*
M-Posttest 20.5				4.0*	7.4*	8.6*	8.8*	9.3*
F Posttest 24.5					3.4*	4.6*	4.8*	5.3*
High Achievement								
M-Pretest 27.9						1.2	1.4	1.9*
M-Posttest 29.1							.2	.7
F-Pretest 29.3								.5
F-Posttest 29.8								

*p <.05

M = Male
F = Female

Table 4. Differences Among Tutor Means: Achievement x Sex x Expectancy

	15.2	21.5	22.2	22.9	28.2	28.9	29.6	29.7
Low Achievement								
MLE 15.2		6.3	7.0	7.7	13.0*	13.7*	14.4*	14.5*
FHE 21.5			.7	1.4	6.7	7.4	8.1*	8.2*
MHE 22.2				.5	6.0	6.7	7.4	7.5
FLE 22.9					5.3	6.0	6.7	6.8
High Achievement								
MLE 28.2						.7	1.4	1.5
MHE 28.9							.7	.8
FLE 29.6								
FHE 29.7								

*p <.05

M = Male
F = Female
HE = High Expectancy
LE = Low Expectancy

Table 5. Source Table for Analysis of Variance: Tutees

Source	DF	Mean Square	F
P	1	1063.142	213.05***
F	1	28.571	<1.00
A	1	14.000	<1.00
E	1	15.017	<1.00
T	1	.260	<1.00
PF	1	13.017	2.60
PA	1	2.160	<1.00
FA	1	236.161	1.95
PE	1	25.786	5.17**
FE	1	591.500	4.88**
AE	1	108.643	<1.00
PT	1	39.447	7.91**
FT	1	85.017	<1.00
AT	1	58.017	<1.00
ET	1	330.286	2.72
PFA	1	.071	<1.00
PFE	1	17.161	3.44
PAE	1	.876	<1.00
FAE	1	54.018	<1.00
PFT	1	23.142	4.64**
PAT	1	1.786	<1.00
FAT	1	68.642	<1.00
PET	1	6.446	1.29
FET	1	4.018	<1.00
AET	1	46.446	<1.00
PFAE	1	14.000	2.81
PFAT	1	42.875	8.59**
PFET	1	.642	<1.00
PAET	1	2.571	<1.00
FAET	1	864.286	7.13**
PFAET	1	24.446	4.90**
Within-Ss Error	96	4.994	
Between-Ss Error	96	121.286	

***p < .001
**p < .01

P = Pretest/Posttest
F = Sex
A = Achievement
E = Expectancy
T = Training

Post hoc analysis of the significant five-way interaction is presented in Table 6. In pursuit of the goals of clarity and interpretability, conceptually meaningful three- and two-way interactions are discussed using the post hoc analysis of the five-way interaction as reference.

Figure 3 illustrates the pretest/posttest by training and sex interaction. Tutees gained significantly more from pretest to posttest when tutored by a trained tutor. This is substantiated by the significant ($p < .01$) training by pretest/posttest interaction, and by the post hoc analysis of the five-way interaction. Cell means comparisons were accomplished by juxtaposing a group on two levels of a variable (e.g., pretest/posttest) and determining if the difference between the cell means equal equaled or exceeded the Tukey honestly significant difference statistic (e.g., 4.44 for the five-way interaction). Six out of eight tests of tutees' pretest to posttest gains with trained tutors were significant ($p < .05$) as compared to three out of eight for tutees who worked with untrained tutors. The primary reason for the variability was tutors' sex. Tutees tutored by trained female tutors scored significantly higher on posttest than on pretest in four out of four cell mean comparisons. For tutees tutored by trained male tutors, only two out of four cell mean comparisons were significant.

Overall, tutees made significant ($p < .001$) gains in their scores on the sight work task from pretest to posttest. The pretest/posttest means were 9.0 and 13.4 respectively.

The pretest/posttest by expectancy interaction was significant ($p < .01$). Reference to Table 6 indicates that for tutees about whom

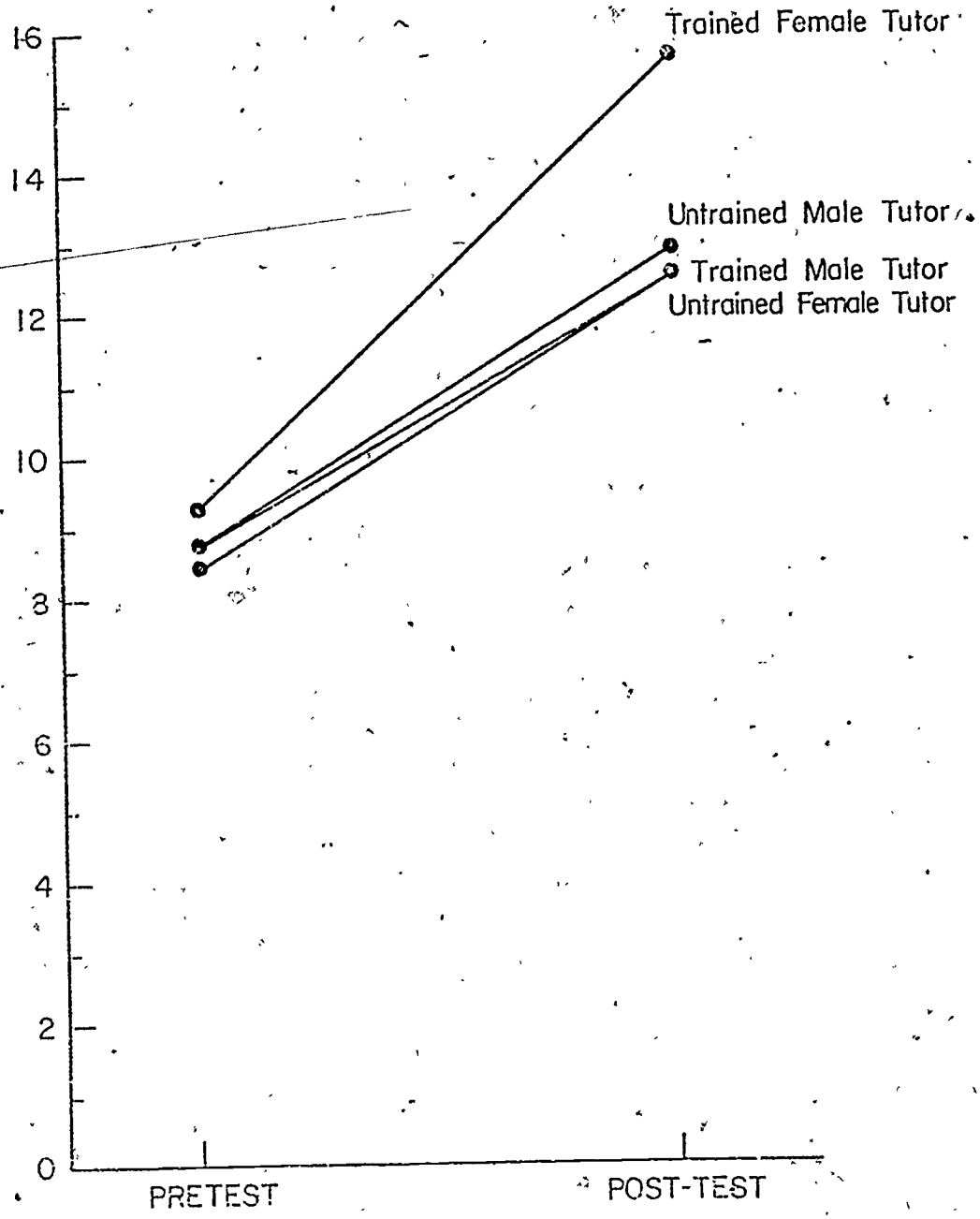


Fig. 3. Tutee Pretest/Posttest Scores x Sex x Training of Tutors

Table 6. Differences Among Tutee Means: Pretest/Posttest x Sex x Achievement x Training x Expectancy

		4.1	5.7	6.1	6.4	7.0	7.3	7.3	7.4	8.3	8.7	9.3	9.9	10.0	10.1	10.
PreMHALET	4.1		1.6	2.0	2.3	2.9	3.2	3.2	3.3	4.2	4.6*	5.2*	5.8*	5.9*	6.0*	6.
PreMHALE	5.7			.4	.7	1.3	1.6	1.6	1.7	2.6	3.0	3.6	4.2	4.3	4.4*	4.
PostMHALET	6.1				.3	.9	1.2	1.2	1.3	2.2	2.6	3.2	3.8	3.9	4.0	4.
PreMLALE	6.4					.6	.9	.9	1.0	1.9	2.3	2.9	3.5	3.6	3.7	3.
PreFLAHE	7.0						.3	.3	.4	1.3	1.7	2.3	2.9	3.0	3.1	3.
PreFLAHET	7.3							0	.1	1.0	1.4	2.0	2.6	2.7	2.8	3.
PreMLAHET	7.3								.1	1.0	1.4	2.0	2.6	2.7	2.8	3.
PreFHALE	7.4									.9	1.3	1.9	2.5	2.6	2.7	2.
PreFHAHET	8.3										.4	1.0	1.6	1.7	1.8	2.
PreFLALET	8.7											.6	1.2	1.3	1.4	1.
PostMLALE	9.3												.6	.7	.8	1.
PreMHAHE	9.9													.1	.2	
PostMHALE	10.0														.1	
PreFLALE	10.1															
PostFHALE	10.3															
PostMLAHET	10.4															
PostFLAHE	10.7															
PreFHAHE	11.0															
PreMHAHET	11.3															
PreFLALET	11.6															
PostFLAHET	11.9															
PostFHAHE	12.4															
PostFHAHET	13.0															
PreFHALET	13.3															
PostMHAHE	13.9															
PostFLALET	15.0															
PreMLAHE	15.0															
PostMHAHET	15.6															
PostFLALE	15.6															
PostFLALET	18.4															
PostMLAHE	18.6															
PostFHALET	23.0															

*p <.05

F = Female
M = Male

HA = High Achievement
LA = Low Achievement
HE = High Expectancy
LE = Low Expectancy

est x

8.3	8.7	9.3	9.9	10.0	10.1	10.3	10.4	10.7	11.0	11.3	11.6	11.9	12.4	13.0	13.3	13.9
4.2	4.6*	5.2*	5.8*	5.9*	6.0*	6.2*	6.3*	6.6*	6.9*	7.2*	7.5*	7.8*	8.3*	8.9*	9.2*	9.8*
2.6	3.0	3.6	4.2	4.3	4.4*	4.6*	4.7*	5.0*	5.3*	5.6*	5.9*	6.2*	6.7*	7.3*	7.6*	8.2*
2.2	2.6	3.2	3.8	3.9	4.0	4.2	4.3	4.6*	4.9*	5.2*	5.5*	5.8*	6.3*	6.9*	7.2*	7.8*
1.9	2.3	2.9	3.5	3.6	3.7	3.9	4.0	4.3	4.7*	5.0*	5.3*	5.6*	6.1*	6.7*	7.0*	7.6*
1.3	1.7	2.3	2.9	3.0	3.1	3.3	3.4	3.7	4.0	4.3	4.6*	4.9*	5.4*	6.0*	6.3*	6.9*
1.0	1.4	2.0	2.6	2.7	2.8	3.0	3.1	3.4	3.7	4.0	4.3	4.6*	5.1*	5.7*	6.0*	6.6*
1.0	1.4	2.0	2.6	2.7	2.8	3.0	3.1	3.4	3.7	4.0	4.3	4.6*	5.1*	5.7*	6.0*	6.6*
.9	1.3	1.9	2.5	2.6	2.7	2.9	3.0	3.3	3.6	3.9	4.2	4.5*	5.0*	5.6*	5.9*	6.5*
	.4	1.0	1.6	1.7	1.8	2.0	2.1	2.4	2.7	3.0	3.3	3.6	4.1	4.7*	5.0*	5.6*
		.6	1.2	1.3	1.4	1.6	1.7	2.0	2.3	2.6	2.9	3.2	3.7	4.3	4.6*	5.2*
			.6	.7	.8	1.0	1.1	1.4	1.7	2.1	2.4	2.7	3.2	3.8	4.1	4.7*
				.1	.2	.4	.5	.8	1.1	1.4	1.7	2.0	2.5	3.1	3.4	4.0
					.1	.3	.4	.7	1.0	1.3	1.6	1.9	2.4	3.0	3.3	3.9
						.2	.3	.6	.9	1.2	1.5	1.8	2.3	2.9	3.2	3.8
							.1	.4	.7	1.0	1.3	1.6	2.1	2.7	3.0	3.6
								.3	.6	.9	1.2	1.5	2.0	2.6	2.9	3.5
									.3	.6	.9	1.2	1.7	2.3	2.6	3.2
										.3	.6	.9	1.4	2.0	2.3	2.9
											.3	.6	1.1	1.7	2.0	2.6
												.3	.8	1.4	1.7	2.3
													.5	1.1	1.4	2.0
														.6	.9	1.5
															.3	.9
																.6

	11.0	11.3	11.6	11.9	12.4	13.0	13.3	13.9	15.0	15.0	15.6	15.6	18.4	18.6	23.0
*	6.9*	7.2*	7.5*	7.8*	8.3*	8.9*	9.2*	9.8*	10.9*	10.9*	11.5*	11.5*	14.3*	15.5*	19.9*
*	5.3*	5.6*	5.9*	6.2*	6.7*	7.3*	7.6*	8.2*	9.3*	9.3*	9.9*	9.9*	12.7*	12.9*	17.3*
*	4.9*	5.2*	5.5*	5.8*	6.3*	6.9*	7.2*	7.8*	8.9*	8.9*	9.5*	9.5*	12.3*	12.5*	16.9*
	4.7*	5.0*	5.3*	5.6*	6.1*	6.7*	7.0*	7.6*	8.7*	8.7*	9.3*	9.3*	12.1*	12.3*	16.7*
7	4.0	4.3	4.6*	4.9*	5.4*	6.0*	6.3*	6.9*	8.0*	8.0*	8.6*	8.6*	11.4*	11.6*	16.0*
	3.7	4.0	4.3	4.6*	5.1*	5.7*	6.0*	6.6*	7.7*	7.7*	8.3*	8.3*	11.1*	11.3*	15.7*
	3.7	4.0	4.3	4.6*	5.1*	5.7*	6.0*	6.6*	7.7*	7.7*	8.3*	8.3*	11.1*	11.3*	15.7*
3	3.6	3.9	4.2	4.5*	5.0*	5.6*	5.9*	6.5*	7.6*	7.6*	8.2*	8.2*	11.0*	11.2*	15.6*
4	2.7	3.0	3.3	3.6	4.1	4.7*	5.0*	5.6*	6.7*	6.7*	7.3*	7.3*	10.1*	10.3*	14.7*
0	2.3	2.6	2.9	3.2	3.7	4.3	4.6*	5.2*	6.3*	6.3*	6.9*	6.9*	9.7*	9.9*	14.3*
4	1.7	2.1	2.4	2.7	3.2	3.8	4.1	4.7*	5.8*	5.8*	6.4*	6.4*	9.2*	9.4*	13.8*
8	1.1	1.4	1.7	2.0	2.5	3.1	3.4	4.0	5.1*	5.1*	5.7*	5.7*	8.5*	8.7*	13.1*
7	1.0	1.3	1.6	1.9	2.4	3.0	3.3	3.9	5.0*	5.0*	5.6*	5.6*	8.4*	8.6*	13.0*
6	.9	1.2	1.5	1.8	2.3	2.9	3.2	3.8	4.9*	4.9*	5.5*	5.5*	8.3*	8.5*	12.9*
4	.7	1.0	1.3	1.6	2.1	2.7	3.0	3.6	4.7*	4.7*	5.3*	5.3*	8.1*	8.3*	12.7*
3	.6	.9	1.2	1.5	2.0	2.6	2.9	3.5	4.6*	4.6*	5.2*	5.2*	8.0*	8.2*	12.6*
.	.3	.6	.9	1.2	1.7	2.3	2.6	3.2	4.3	4.3	4.9*	4.9*	7.7*	7.9*	12.3*
		.3	.6	.9	1.4	2.0	2.3	2.9	4.0	4.0	4.6*	4.6*	7.4*	7.6*	12.0*
			.3	.6	1.1	1.7	2.0	2.6	3.7	3.7	4.3	4.3	7.1*	7.3*	11.7*
				.3	.8	1.4	1.7	2.3	3.4	3.4	4.0	4.0	6.8*	7.0*	11.4*
					.5	1.1	1.4	2.0	3.1	3.1	3.7	3.7	6.5*	6.7*	11.1*
						.6	.9	1.5	2.6	2.6	2.9	2.9	5.7*	5.9*	10.3*
							.3	.9	2.0	2.0	2.6	2.6	5.4*	5.6*	10.0*
								.6	1.7	1.7	2.3	2.3	5.1*	5.3*	9.7*
									1.1	1.1	1.7	1.7	4.5*	4.7*	9.1*
										0	.6	.6	3.4	3.6	8.0*
											.6	0	3.4	3.6	8.0*
													2.8	3.0	7.4*
													2.8	3.0	7.4*
														.2	4.6*
															4.4*

tutors had high expectancies, only three of eight pretest to posttest gains were significant, compared to six of eight significant pretest to posttest gains for tutees about whom tutors had low expectancies.

Figure 4 illustrates the same data, revealing that a significant (see Table 7) pretest difference in pretest scores for tutees about whom tutors had high or low expectancies is not significant on the posttest.

Table 7. Differences Among Tutee Means: Pretest/Posttest x Expectancy

		8.4	9.6	13.3	13.5
Pretest					
Low Expectancy	8.4		1.2*	4.9*	5.1*
High Expectancy	9.6			3.7*	3.9*
Posttest					
Low Expectancy	13.3				.2
High Expectancy	13.5				

*$p < .05$

The sex by achievement by expectancy by training interaction and the sex by expectancy interaction fail to yield relevant data. Figure 5 illustrates the two-way interaction. Table 6 results indicate that tutees tutored by untrained, low achievement, high expectancy males got higher scores on the pretest and the posttest than their tutee counterparts tutored by untrained, low achievement, high expectancy females.

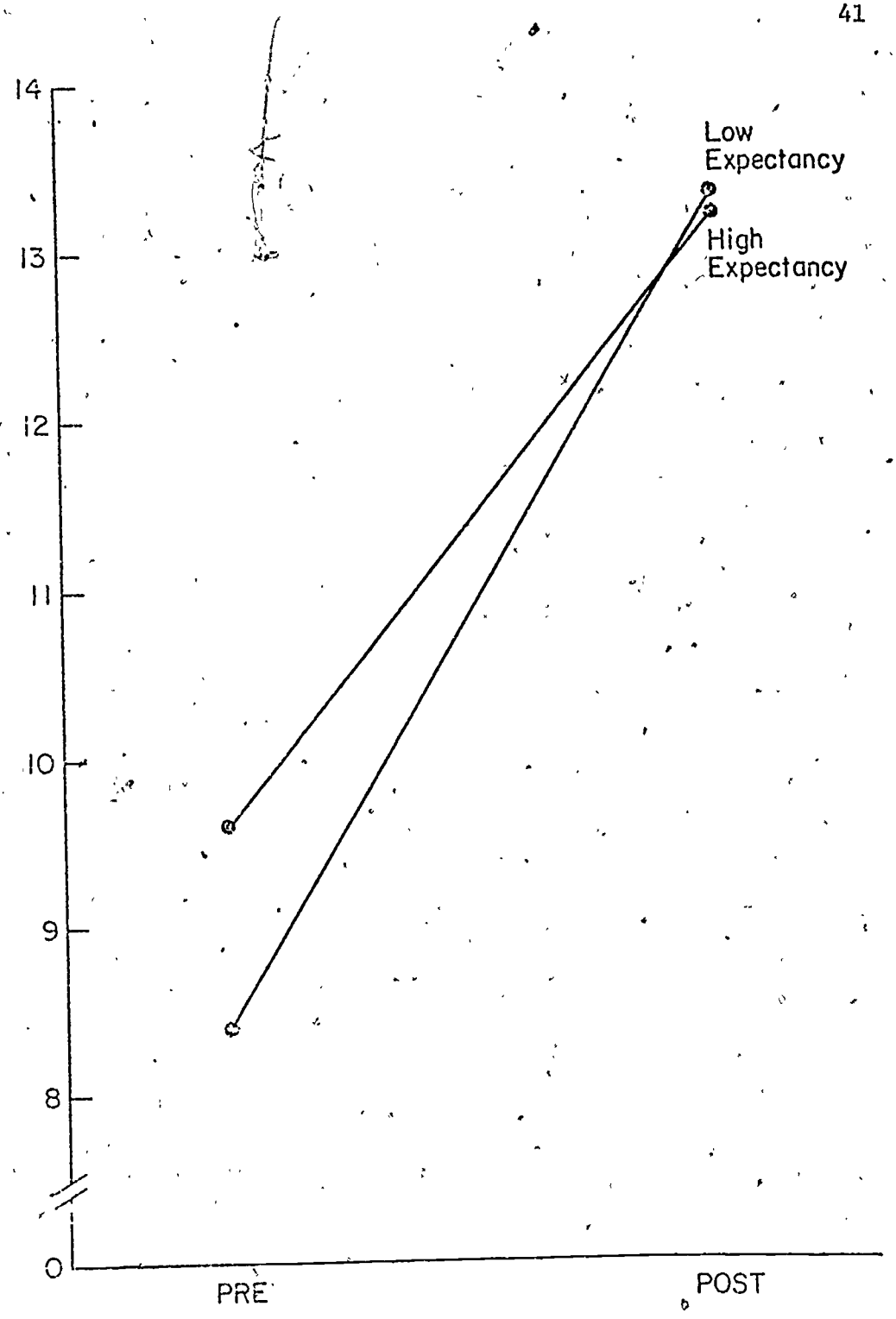


Fig. 4. Tutee Scores x Pretest/Posttest x Expectancy

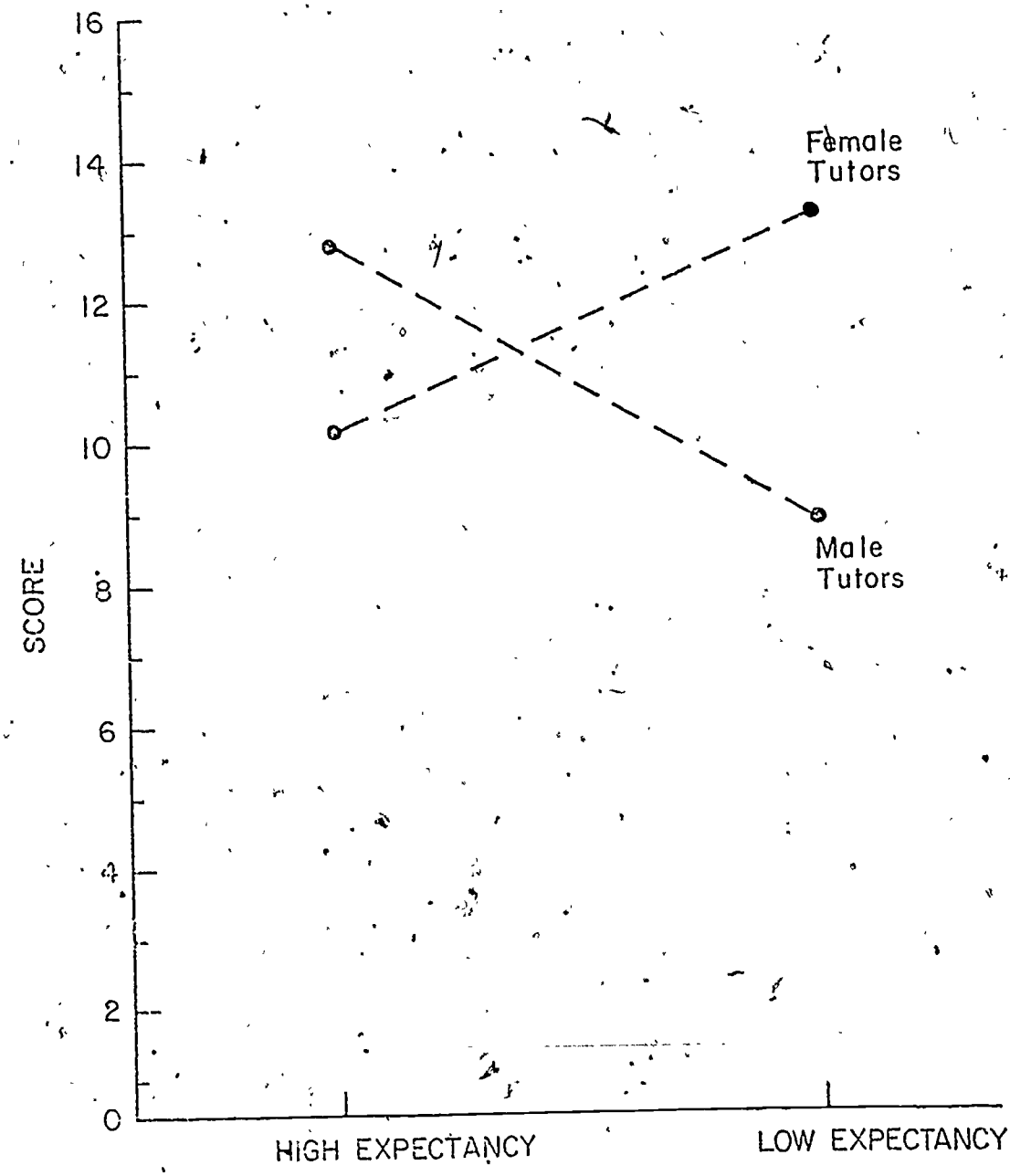


Fig. 5. Tutor Scores x Sex x Expectancy

In a similar confusing fashion, tutees tutored by trained, high achievement, low expectancy females achieved higher scores on the pretest and the posttest than did their tutee counterparts tutored by trained, high achievement, low expectancy males. In summary, two of eight mean cell comparisons for the low expectancy group were significant. The size of these differences within the two levels of achievement were sufficiently large to yield a significant interaction.

Observation of the tutor during the tutoring session provided data on the following variables:

1. number of cards presented to tutee
2. frequency of negative behaviors
3. frequency of providing or accepting a word incorrectly
4. frequency of reinforcement to tutee
5. frequency of providing the word as a form of feedback
6. frequency of showing the picture as a form of feedback
7. frequency of re-presenting the card to tutee after feedback

These variables were analyzed by seven four-way analyses of variance with sex, training of tutor, expectancy, and tutor achievement as the between-subjects variables.

The first two dependent variables analyzed from the observation data (number of cards presented and frequency of negative behaviors) failed to yield significant results.

Table 8 replicates the source table for the analysis of variance of providing or accepting the incorrect word to or from the tutee.

Table 8. Analysis of Variance: Providing Incorrect Word to Tutee

Source	DF	Mean Square	F
F	1	.893	<1.00
A	1	34.321	13.50**
E	1	1.750	<1.00
T	1	1.750	<1.00
FA	1	3.571	1.40
FE	1	.571	<1.00
AE	1	3.571	1.40
FT	1	3.571	1.40
AT	1	5.143	2.02
ET	1	2.286	<1.00
FAE	1	2.893	1.14
FAT	1	.893	<1.00
FET	1	.036	<1.00
AET	1	4.321	1.69
FAET	1	.143	<1.00
Between-Ss Error	96	2.55	

**p <.01

F = Sex
A = Achievement
E = Expectancy
T = Training

High achieving tutors provided or accepted a word incorrectly significantly less ($\bar{x}$ = .23) than did low achieving tutors ($\bar{x}$ = 1.34). The issue of significance and meaningfulness will be discussed in Chapter 5 in relation to this result.

The analysis of frequency of tutor reinforcement to the tutee is summarized in Table 9. This frequency was computed as a ratio of the number of reinforcements to tutee over the total number of cards presented. The resulting figure provides a measurement of how many times the tutor reinforced the tutee out of the total opportunities for such reinforcement.

Table 9. Analysis of Variance: Reinforcement to Tutee

Source	DF	Mean Square	F
F	1	24.456	<1.00
A	1	217.096	2.79
E	1	1.538	<1.00
T	1	4277.150	55.01**
FA	1	129.948	1.67
FE	1	20.116	<1.00
AE	1	5.902	<1.00
FT	1	578.163	7.44**
AT	1	141.361	1.82
ET	1	37.559	<1.00
FAE	1	278.103	3.58
FAT	1	51.044	<1.00
FET	1	.004	<1.00
AET	1	54.516	<1.00
FAET	1	358.550	4.61**
Between-Ss Error	96	77.547	

**p <.01

F = Sex
A = Achievement
E = Expectancy
T = Training

The ratio of reinforcement to the tutee by sex, training, achievement, and expectancy are illustrated in Figure 6. Table 10 delineates the results of Tukey's honestly significant difference post hoc test. As seen in the graph, the overall tendency is for trained tutors to provide more reinforcement to tutees than untrained tutors. Specifically, high achievement, trained female tutors of either high or low expectancy gave more reinforcement to tutees than did their untrained counterparts. This difference approximated significance (within .003 of a point) for low achievement, high expectancy, trained female tutors as compared to low achievement, high expectancy, untrained female tutors.

The sex by training significant interaction ($p < .01$) derives from the greater frequency of reinforcement from trained female tutors. Two of four (nearly three of four) cell mean comparisons yield significantly ($p < .05$) higher frequency of reinforcement by trained female tutors as compared to untrained female tutors. A similar comparison of trained versus untrained male tutors yields no significant cell mean comparisons.

Overall, training was a significant ($p < .01$) variable with a mean of .506 ratio of reinforcement from trained tutors and a mean of .116 ratio of reinforcement for untrained tutors.

The ratio of the number of times the tutor provided the word for the tutee over the total number of times feedback was given was analyzed by an analysis of variance. Table 11 summarizes the resulting source table.

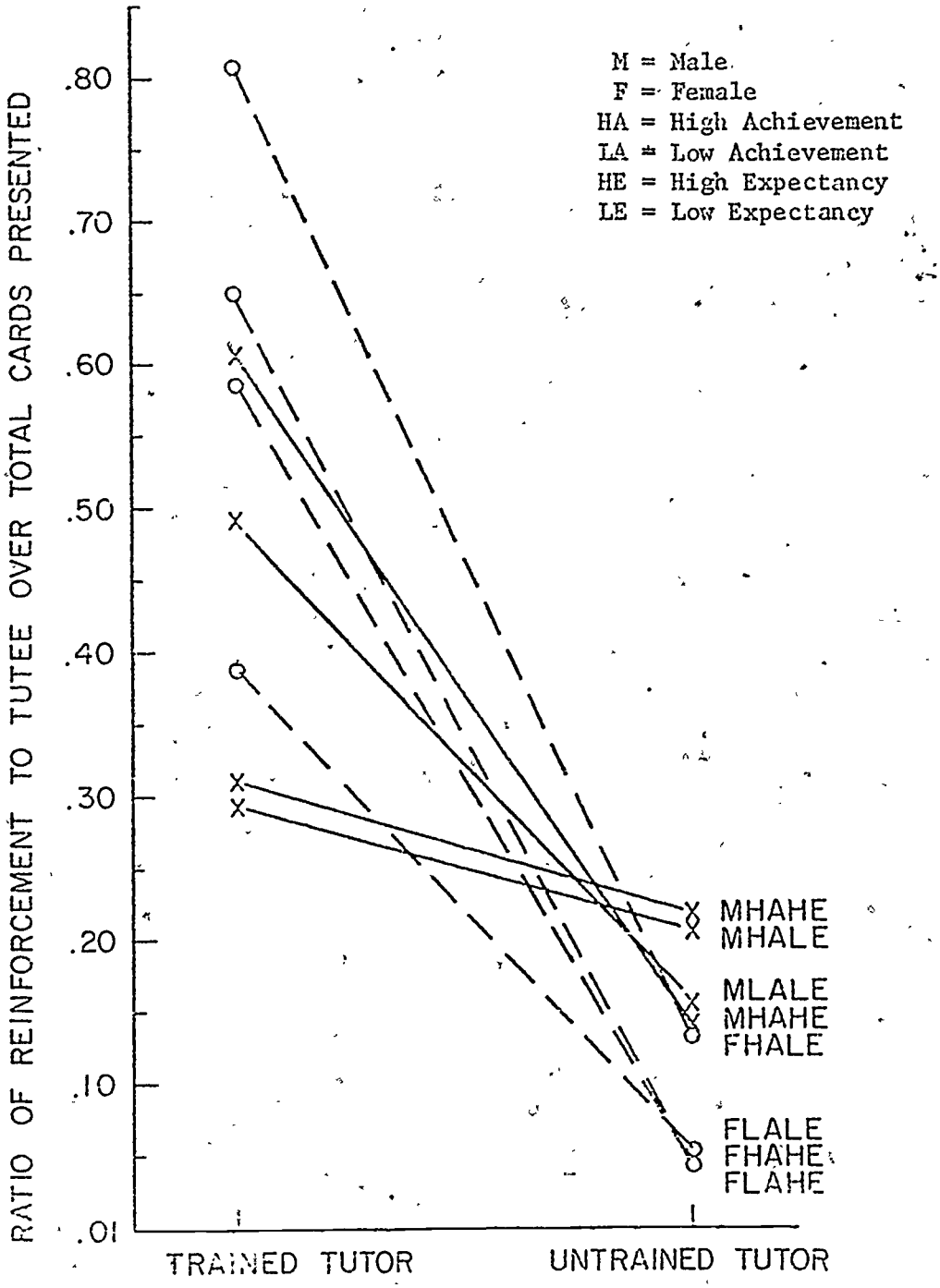


Fig. 6. Ratio of Reinforcement x Training x Sex x Achievement x Expectancy of Tutor

Table 10. Differences Among Behavioral Data Means: Ratio of Reinforcement

	UNTRAINED								TRAINED							
	.029	.032	.043	.130	.131	.142	.206	.212	.287	.310	.358	.481	.558	.601	.645	.811
FHAHE .029		.003	.014	.101	.102	.113	.177	.183	.264	.287	.335	.458	.535*	.578*	.622*	.788*
FLAHE .032			.011	.098	.099	.110	.174	.180	.261	.284	.332	.455	.532	.575*	.619*	.785*
FLALE .043				.087	.088	.099	.163	.169	.250	.273	.321	.444	.521	.564*	.608*	.774*
FHALE .130					.001	.012	.076	.082	.163	.186	.234	.357	.434	.477	.521	.687*
MHAHE .131						.011	.075	.081	.162	.185	.233	.356	.433	.476	.520	.686*
MLALE .142							.064	.070	.151	.174	.222	.345	.422	.465	.509	.675*
MHALE .206								.006	.087	.110	.168	.291	.368	.411	.455	.621*
MLAHE .212									.081	.104	.152	.275	.352	.395	.439	.605*
MHALE .287										.023	.071	.194	.271	.314	.358	.524
MLAHE .310											.048	.171	.248	.291	.335	.501
FLALE .358												.123	.200	.243	.287	.453
MLALE .481													.077	.120	.164	.330
FLAHE .558														.043	.087	.253
MHAHE .601															.044	.210
FHAHE .645																.166
FHALE .811																

*$p < .05$

F = Female
M = Male

HA = High Achievement
LA = Low Achievement

HE = High Expectancy
LE = Low Expectancy

Table 11. Analysis of Variance: Providing Word as Feedback

Source	DF	Mean Square	F
F	1	159.305	<1.00
A	1	.201	<1.00
E	1	72.930	<1.00
T	1	2453.248	14.09**
FA	1	201.451	1.16
FE	1	96.879	<1.00
AE	1	2.527	<1.00
FT	1	113.412	<1.00
AT	1	.424	<1.00
ET	1	150.822	<1.00
FAE	1	6.604	<1.00
FAT	1	25.380	<1.00
FET	1	6.573	<1.00
AET	1	127.305	<1.00
FAET	1	5.212	<1.00
Between-Ss Error	96	174.167	

**p <.001

F = Sex
A = Achievement
E = Expectancy
T = Training

Tutors who had been trained provided the word for the tutee when they gave feedback significantly ($p < .01$) more often than did tutors who were not trained. The mean ratio for the trained tutors was .861 compared to the mean ratio of .565 for untrained tutors.

The frequency of tutor showing the picture to the tutee as the form of feedback was analyzed in the same manner as providing the word as feedback. A ratio was formed by placing the frequency of tutor

showing the picture to tutee over the total number of times tutor provided feedback. Table 12 presents the source table on this analysis of variance.

The results of the post hoc test of the significant ($p < .01$) disordinal achievement by training interaction is presented in Table 13. Figure 7 illustrates this interaction.

Low achievement trained tutors showed the picture to the tutee as a form of feedback significantly more often than did high or low achievement untrained tutors. High achievement trained tutors also used the picture as a form of feedback significantly more often than did either high or low achievement tutors. The training effect was significant ($p < .01$) with a mean of .752 for trained tutors compared to a mean of .300 for untrained tutors.

The final variable recorded on the tutor behavior observation schedule was the frequency of re-presenting the card (in essence, re-asking the question) to the tutee following feedback. As with the preceding two variables, this dependent variable was computed as a ratio of the frequency of re-presenting the card over the total frequency of feedback. Table 14 summarizes the source table for this analysis of variance.

A post hoc analysis of the sex by achievement by expectancy interaction failed to reveal any significant cell mean comparisons. The Tukey honestly significant difference test is obviously a more stringent test than the F-ratio. The post hoc analysis of the significant sex by

Table 12. Analysis of Variance: Showing Picture as Feedback

Source	DF	Mean Square	F
F	1	287.348	1.747
A	1	10.511	<1.00
E	1	.131	<1.00
T	1	5739.965	34.90**
FA	1	125.625	<1.00
FE	1	85.748	<1.00
AE	1	538.073	3.27
FT	1	91.143	<1.00
AT	1	796.163	4.84**
ET	1	149.723	<1.00
FAE	1	6.436	<1.00
FAT	1	180.723	1.10
FET	1	343.325	2.09
AET	1	73.800	<1.00
FAET	1	167.633	1.02
Between-Ss Error	96	164.466	

**p <.01

F = Sex
A = Achievement
E = Expectancy
T = Training

Table 13. Differences Among Behavioral Data Means: Picture as Feedback

		Untrained		Trained	
		.206	.394	.678	.827
High Achievement	.206		.188	.472*	.621*
Low Achievement	.394			.284*	.433*
Low Achievement	.678				.149
High Achievement	.827				

*p <.05

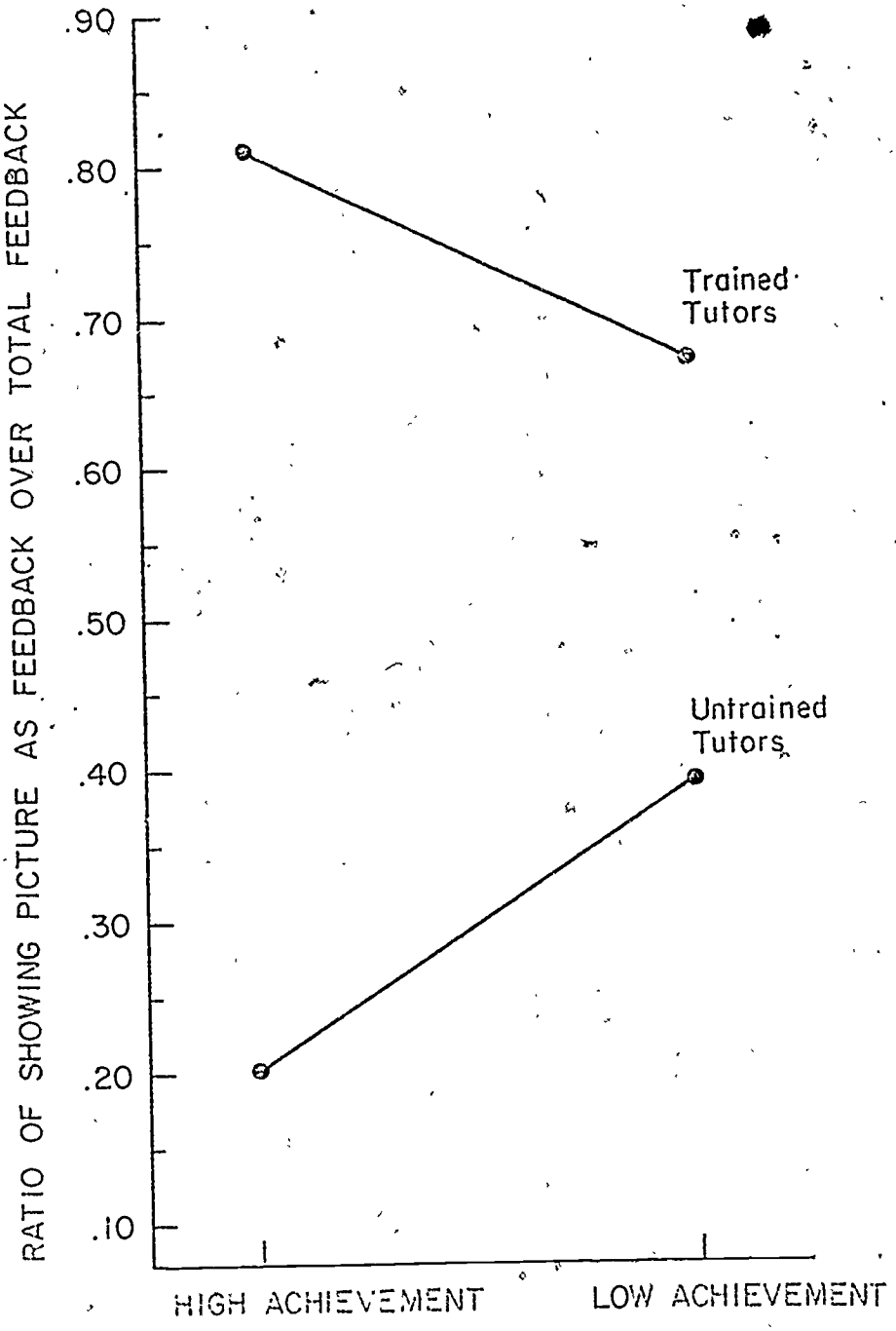


Fig. 7. Showing Picture as Feedback x Achievement Level x Training

Table 14. Analysis of Variance: Re-presenting Card to Tutee

Source	DF	Mean Square	F
F	1	58.514	<1.00
A	1	317.370	3.11
E	1	40.508	<1.00
T	1	9412.241	92.32**
FA	1	693.316	6.80**
FE	1	52.289	<1.00
AE	1	58.972	<1.00
FT	1	3.192	<1.00
AT	1	6.542	<1.00
ET	1	43.845	<1.00
FAE	1	443.521	4.35**
FAT	1	325.944	3.20
FET	1	124.756	1.22
AET	1	270.089	2.65
FAET	1	252.510	2.48
Between-Ss Error	96	101.948	

**p <.01

F = Sex
A = Achievement
E = Expectancy
T = Training

achievement interaction is presented in Table 15. High achieving female tutors re-presented the cards to the tutees significantly more often than did low achieving female tutors. Training of tutors was a significant main effect, with trained tutors re-presenting the card to the tutees significantly more frequently than did untrained tutors, with means of .658 and .078 respectively.

Table 15. Differences Among Means: Re-presenting Card to Tutee

	.259	.320	.371	.523
FLA .259		.061	.112	.264*
MHA .320			.051	.203
MLA .371				.152
FHA .523				

*p < .05

M = Male
F = Female

LA = Low Achievement
HA = High Achievement

CHAPTER 5

DISCUSSION

Results of this study indicate that both tutors and tutees derive academic benefit from a brief (e.g., ten minutes) peer tutoring session. This positive effect was more pronounced for low achieving than for high achieving tutors. Low achieving tutors had significantly lower pretest and posttest scores than did high achieving tutors. However, both male and female low achieving tutors made significant gains from pretest to posttest. Gains for high achieving tutors are rendered uninvestigable because of the ceiling effect. High achieving tutors increased from a pretest mean of 28.7 to a posttest mean of 29.5 out of a maximum total of 30.

Training of tutors significantly increased the pretest to posttest gains for both tutors and tutees. Possible reasons for these data lie in the behavioral data recorded during the tutoring session. Trained tutors provided more reinforcement, more word and/or picture feedback, and re-asked the question of the tutees more often than did untrained tutors. The provision of feedback necessitated rehearsal of the correct response by the tutors. Re-presenting the card after feedback resulted in rehearsal for both tutees and tutors. Trained female tutors provided more reinforcement and re-presented the card more often than did trained male tutors. This result has limited practical value,

however. It would be an unfortunate result for teachers to deprive males of the benefits of the tutoring role because of a significant sex effect in research.

Overall, expectancy did not have an effect on tutee pretest to posttest gains nor on tutor behavior during the tutoring session. Tutees about whom tutors had low expectancies gained more significantly from pretest to posttest than did tutees about whom tutors had high expectancies. However, low expectancy tutees were significantly lower on pretest than were high expectancy tutees. There were no significant trends in tutor behavior that can be attributed to the expectancy-biasing instructions.

Achievement of tutors did not yield a significant difference in gains realized by tutees. Tutees tutored by trained or untrained low achieving tutors evidenced similar pretest to posttest gains, as did tutees tutored by trained or untrained high achieving tutors. Tutor achievement level was a significant variable in the provision or acceptance of a word incorrectly to or from a tutee. High achieving tutors were significantly more accurate than low achieving tutors (means of 1.34 and .23, respectively). This statistically significant difference is less meaningful when it is known that the low achieving tutor, on the average, was inaccurate one and one-third times out of seventy cards, and that this inaccuracy had no significant effect on the tutees' gains from the learning situation. The use of self-correcting materials in an applied situation generally allows the tutor achievement level to be irrelevant.

Failure to reject the null hypothesis is moot with regard to statistical inferences. However, in applied research, finding that a variable (e.g., tutor achievement level) fails to differentiate groups can have meaningful implications. In this case, there has been a common sense based assumption that high achieving tutors were the only effective tutors. The fact that tutee gains were not dependent on tutor achievement level yields the joyous embrace (versus mere acceptance) of the null hypothesis in a practical sense, not a statistical inferential sense.

In light of this, Chapter 2 delineated nine hypotheses with the predicted direction of effects reflecting the viewpoint of the author. This viewpoint was based on both classroom experience and on the literature review. Of these hypotheses, two were not supported by the data. First, hypothesis 4 stated that tutor reinforcement training would not result in significantly higher tutor gain scores. Tutor training did result in significant pretest to posttest gains for the tutors, as well as for the tutees. This result further confirms the positive value of the tutoring role and of a brief amount of corrective feedback training.

The second unsupported hypothesis is that there would be no significant sex effect. The most effective peer tutors as measured by tutee pretest to posttest gains and by frequency of various corrective feedback procedures were trained females, as compared to trained males or to untrained male or female tutors. This significant result perhaps reflects socialization differences between the sexes, but it has little practical value in making peer tutoring recommendations to teachers.

From an applied stance, the following recommendations have been made (Conrad, 1975) to teachers on the basis of this research.

A. A minimum amount of training on corrective feedback procedures can increase tutor effectiveness and both tutor and tutee gains.

B. The amount of attention and stimulation provided by peers may not be influenced by expectations about tutee performance.

C. Tutor training and the use of self-correcting materials can yield effective tutors regardless of ability level.

Future Research

A re-analysis of the data from this study analyzing tutee and tutor pretest to posttest gains on the basis of which tutors provided feedback, regardless of the etiology of their skill (training or modeling from a teacher) would yield further information on the relevance of the tutor behavior to achievement.

An implicit assumption of open classrooms and of tutor training programs is that peer tutoring is trans-situational, i.e., present in academic and play situations. The author has pilot-tested an observation instrument and plans to conduct a study to investigate whether (a) peer tutoring is naturalistically observed in a variety of situations, (b) peer tutoring skills (after tutor training) will generalize to a variety of situations, and (c) peer tutoring skills are observed some time (e.g., three months) after training.

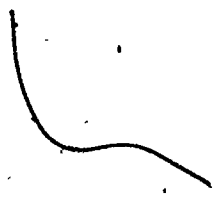
A compendium of self-correcting tasks amenable to peer tutoring situations and various techniques of implementing peer tutoring

into daily classroom routine is currently being compiled by the author with the help of her friends: teachers and administrators in Lincoln, Nebraska.

The extent of the effect of peer tutors' expectations could be more realistically studied by observing peer tutors tutoring children about whom they indicated they had low expectations versus children about whom they indicated they had high expectations. This procedure of capitalizing on naturalistically-formed expectations would remove the doubt as to whether peers do not teach differentially based on expectations, or whether the instructions given to tutors in this study were not sufficient to generate expectations.

APPENDIX A

TUTOR BEHAVIOR OBSERVATION FORM



Tutor Behavior Observation Form

Tutor_____

Tutee_____

Sex F M
Achievement H L
Expectancy H L
Training Y N
School _____
Teacher _____

Presents card.

R+

R+

R+

Provides correct word or picture if incorrect or no answer.

word
picture

word
picture

word
picture

Repeats a question or re-presents card if incorrect answer is given.

Punative words or gestures.

Repeats word as acknowledgement.

Says word incorrectly.

REFERENCES

Amster, H., Davis, B. G., and Logan, D. D. Effectiveness of Schell's peer tutoring program in mathematics for junior high school students. Paper presented at the annual conference of the American Educational Research Association, Washington, D.C., March 1975.

Anderson, D., and Rosenthal, R. Some effects of interpersonal expectancy and social interaction on institutionalized retarded children. Proceedings of the 76th Annual Convention of the American Psychological Association, 1968, 3, 479-480.

Barth, R. S. Open education and the American school. New York: Agathon Press, 1972.

Bausell, R. B., Moody, W. B., and Walzl, F. N. A factorial study of tutoring versus classroom instruction. American Educational Research Journal, 1972, 9, 591-599.

Beez, W. V. Influence of biased psychological reports on teacher behavior and pupil performance. In M. B. Miles and W. W. Charters, Jr. (Eds.), Learning in social settings. Boston: Allyn and Bacon, Inc., 1970.

Bernard, H. W. Psychology of learning and teaching. New York: McGraw Hill, 1972.

Brophy, J. E., and Good, T. L. Teacher-student relationships: Causes and consequences. New York: Holt, Rinehart & Winston, 1974.

Brown, J. C. Effects of token reinforcement adminstered by peer-tutors on pupil reading achievement and tutor collateral behavior. Dissertation Abstracts International, 1972, 32, 3775.

Bruner, J. S. Needed: A theory of instruction. In S. Dropkin, H. Full, and E. Schwarcz (Eds.), Contemporary American Education. New York: Macmillan Co., 1965, 355-364.

Claiborn, W. L. Expectancy effects in the classroom: A failure to replicate. Journal of Educational Psychology, 1969, 60, 377-383.

Cloward, R. D. Studies in tutoring. Journal of Experimental Education, 1967, 36, 14-25.

Conn, L., Edwards, C., Rosenthal, R., and Crowne, D. Perceptions of emotion and response to teachers' expectancy by elementary school children. Psychological Reports, 1968, 22, 27-34.

Conrad, E. Peer tutoring: A cooperative learning experience. Tucson: Arizona Center for Educational Research and Development, University of Arizona, 1975.

Durlack, J. A. Ninth graders as student aides: Making use of the helper therapy principle. Psychology in the Schools, 1973, 3, 334-339.

Ellson, D. G., Barber, L., Engle, T. L., and Kampworth, L. Programmed tutoring: A teaching aid and a research tool. Reading Research Quarterly, 1965, 1, 77-127.

Evans, J., and Rosenthal, R. Interpersonal self-fulfilling prophecies: Further exprapolations from the laboratory to the classroom. Proceedings of the 77th Annual Convention of the American Psychological Association, 1969, 4(1), 371-372.

Feldman, R. S., and Allen, V. L. Learning through tutoring: Low-achieving subjects as tutors. Educational Document Reproduction Service, ED 066 667, Washington, D.C., May, 1972.

Fine, L. The effects of positive teacher expectancy on the reading achievement and IQ gains of pupils in grade two. Dissertation Abstracts International, 1972, 33(A), 1510-1511.

Fleming, E. S., and Anttonen, R. Teacher expectancy as related to the academic and personal growth of primary-age children. Monographs of the Society for Research in Child Development, 1971, 36(5).

Greenwood, C. R. Effects of elementary aged, peer, behavior managers on small group, written academic response rates. Dissertation Abstracts International, 1973, 33, 3381.

Grieger, R. M. Pygmalion revisited: A loud call for caution. Interchange, 1971, 2(4), 78-91.

Guarnaccia, V. J. Pupil-tutoring in elementary school aritlmetic instruction. Dissertation Abstracts International, 1973, 33 3283.

Hagen, J. W., and Moeller, T. Cross-age tutoring. Educational Document Reproduction Service, ED 085 090, Washington, D.C., July, 1971.

Hamblin, J. A., and Hamblin, R. L. On teaching disadvantaged pre-schoolers to read: A successful experiment. American Educational Research Journal, 1972, 9, 209-216.

Hamblin, R. L., Buckholdt, D., Ferritor, D., Kazloff, M., and Blackwell, L. The humanization process. New York: John Wiley & Sons, 1971.

Hunt, B. B. Some effects of peer tutoring on self-esteem and racial attitude. Dissertation Abstracts International, 1971, 32, 92.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks/Cole Publishing Company, 1968.

Lippitt, P., Eiseman, J., & Lippitt, R. Children can teach other children. The Instructor, 1969, 9, 46-52.

Lippitt, P., and Lohman, J. E. A neglected source: Cross-age relationships. Children, 1965, 12, 113-117.

Mason, E. J. Teacher's observations and expectations of boys and girls as influenced by biased psychological reports and knowledge of the effects of bias. Dissertation Abstracts International, 1973, 33(4-A).

Medinnus, G., and Unruh, R. Teacher expectations and verbal communication. Paper presented at the annual meeting of the Western Psychological Association, Denver, April, 1971.

Mollod, R. W. Pupil-tutoring as a part of reading instructions in the elementary school. Dissertation Abstracts International, 1970, 31, 2260.

Morgan, R. F., and Toy, T. B. Learning by teaching: A student-to-student compensatory tutoring program in a rural school system and its relevance to the educational cooperative. Psychological Record, 1970, 20, 159-169.

Murray, B. B. Individualized amelioration of learning disability through parent-helper-pupil involvement. Dissertation Abstracts International, 1972, 33, 1554.

Myers, K. E., Travers, R. M. W., and Sanford, M. E. Learning and reinforcement in student pairs. Journal of Educational Psychology, 1965, 56, 67-72.

Niedermyer, F. C., and Ellis, P. A. The development of a tutorial program for kindergarten reading instruction. San Francisco: Southwest Regional Laboratory for Educational Research and Development, 1970.

Pellegrini, R., and Hicks, R. Prophecy effects and tutorial instruction for the disadvantaged child. American Education Research Journal, 1972, 9, 413-419.

Resnick, L. B. Open education: Some tasks for technology. Educational Technology, 1972, 12, 70-76.

Rosenthal, R., and Jacobson, L. Teacher expectancies: Determinants of pupils' IQ gains. Psychological Reports, 1966, 19, 115-118.

Rubovits, R., and Maehr, M. Pygmalion analyzed: Toward an explanation of the Rosenthal-Jacobson findings. Journal of Personality and Social Psychology, 1971, 19, 197-203.

Shaver, J. P., and Nuhn, D. The effectiveness of tutoring underachievers in reading and writing. Journal of Educational Research, 1971, 65, 107-112.

Snapp, M., Oakland, T., and Williams, F. C. A study of individualizing instruction by using elementary school children as tutors. Journal of School Psychology, 1972, 10, 1-8.

Snow, R. E. Unfinished Pygmalion. Contemporary Psychology, 1969, 14, 197-199.

Stainback, W. C. Effects of pupil-to-pupil tutoring on arithmetic achievement and personal and social adjustment. Dissertation Abstracts International, 1972, 32, 4456.

Tannenbaum, A. J. An evaluation of STAR: A non-professional tutoring program. Teachers College Record, 1968, 69, 433-448.

Tanner, L. N., and Lindgren, H. C. Classroom teaching and learning. New York: Holt, Rinehart, & Winston, 1971.

Thorndike, R. L. Review of Pygmalion in the classroom. American Education Research Journal, 1968, 708-711.

Torrance, E. P. Peer influences on preschool children's willingness to try difficult tasks. Journal of Psychology, 1969, 72, 189-194.

Von Harrison, G., Nelson, W., and Tregaskis, L. The use of a structured tutorial reading program in teaching nonreading second graders in Title I schools to read. Paper presented at the American Educational Research Association's Annual Meeting, Chicago, April, 1972.

Wolfson, B. J. Pupil and teacher roles in individualized instruction. In M. L. Silberman, J. S. Allender, and J. M. Yanoff (Eds.), The psychology of open teaching and learning: An inquiry approach. Boston: Little, Brown & Co., 1972.